Evaluation of *Understanding Society*: Impact and views of data users

Report by CFE Research for ESRC

May 2022

Dr Sarah Harding
Dr Sarah Tazzyman
John Higton
Professor Jennifer Roberts
Professor Nigel Rice

CFE research

For more information about this report please contact John Higton:
CFE Research, Phoenix Yard, Upper Brown Street, Leicester, LE1 5TE
T: 0116 229 3300 john.higton@cfe.org.uk www.cfe.org.uk

Established since 1997, CFE is an independent not-for-profit company specialising in the provision of research and evaluation services across a broad field of education, employment and skills.

Contents

Glossary

Bias	Some members of the intended population have a higher (or lower) probability of being sampled than others
BHPS	British Household Panel Survey (precursor to UKHLS, and the waves in BHPS date back to 1991)
Cross-sectional	Approach to data collection and analysis that takes the views of the sample in any wave/year in isolation
DVLA	Driver and Vehicle Licensing Agency
DWP	Department for Work & Pensions
ESRC	Economic and Social Research Council
Green decision tree	A 'tree-like model' of decisions to visually and explicitly represent decisions and decision-making. Used in machine learning to make predictions of target variables
HE	Higher education
IFS	Institute for Fiscal Studies
IT	Information technology
Longitudinal	Approach to data collection and analysis that samples the same people across waves, resulting in repeated measures for the same individual over time
LSOA	Lower-layer Super Output Area
Non-response	Participants do not respond to a survey
Non-response bias	Bias that occurs when people are unwilling/unable to respond to a survey have different observed or unobserved characteristics or factors from those who have responded
NPD	National Pupil Database
ONS	Office for National Statistics
PhD	Doctor of Philosophy (highest academic level qualification)
PPU	Policy and Partnerships Unit
REF	Research Excellence Framework
UKDS	UK Data Service
UKHLS	UK Household Longitudinal Study

Executive Summary

- *Understanding Society* has a global reputation, and users think that the UK data infrastructure would be significantly weaker without it. Users are overwhelmingly positive about *Understanding Society* and often see it as a 'go-to' dataset. The perceived value of the dataset increases with usage.

- The most valued aspects of *Understanding Society* cited by users include the quality and timeliness of the data, the number and richness of variables available, the large sample size that is representative of the whole UK population, and the longitudinal design that follows both individuals and households. This allows researchers to explore multiple dimensions of the same person's life over time.

> "I believe that *Understanding Society* is the envy of the world. When I have presented at international conferences, colleagues have come up and said, 'Wow, we wish we would have something like that in our country." *Academic user* – ***Work, education, skills & technology***

- *Understanding Society* has led to a range of benefits and impacts, including academic, social, economic and methodological definitions of impact. Twelve case studies included in this report document the pathways to impact and different ways that impact can be achieved.

> "*Understanding Society* is a great national treasure. We should celebrate it and not be shy about it, boast about it, tell the world, and let the world use it." - ***Academic user (focus group)– Society, governance & security***

- Tracking policy impact can be difficult, but research and findings from using *Understanding Society* feed into the evidence base which is then considered at ministerial and policy level.

- The quality of outputs, contribution to knowledge across a wide range of disciplines and the sustained longitudinal data collection are key elements that support continued funding of *Understanding Society*.

- Changes to the dataset, such as smaller population sizes or fewer variables would adversely compromise the research quality and subsequent impacts that could be achieved through analysis.

Understanding Society, the UK Household Longitudinal Study, is the largest longitudinal household panel study of its kind and one of the UK's leading data collections. The ESRC has committed over £160 million to the study, the greatest amount of public funding ESRC have ever made to a single research investment. Supporting social, economic and academic research, the study aims to track changes in the lives of people, and in their homes and communities.

Understanding Society data is collected annually from approximately 25,000 households across all four UK nations. The study explores six core domains (education and training; employment; family, household and social connectedness; civic engagement; income, consumption and wealth; and health and health behaviours). Fieldwork was initiated in 2009 as the successor to the British Household Panel Survey (BHPS) and, at the date of writing, consists of 11 completed data waves.

"I hope [*Understanding Society*] continues. I hope it doesn't get thinned out or anything because it's a very valuable resource that's good for the economy, good for the country, and good for research." ***Academic user – Society, governance & security***

The ESRC commissioned CFE Research to undertake an independent evaluation of *Understanding Society* to inform future business cases and funding decisions. This report will complement ongoing benefits realisation reports and case study reports that document the activities and impact associated with *Understanding Society*.

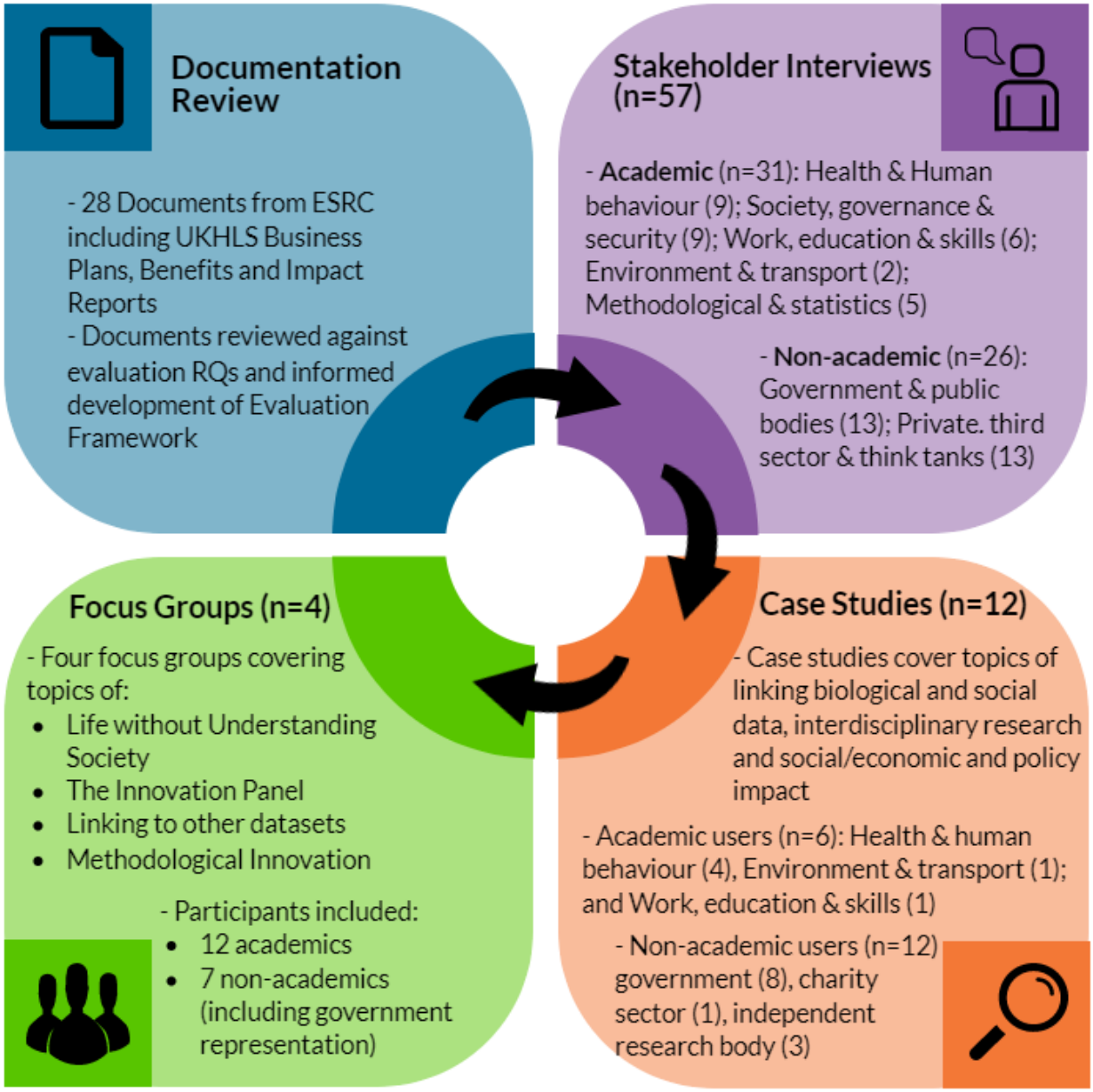
The evaluation addresses five research questions to gauge how the use and impact of *Understanding Society* has changed over time:

1) What do stakeholders think of *Understanding Society*?
2) How is *Understanding Society* used by stakeholders?
3) What have stakeholders been able to achieve which would have been difficult or impossible if *Understanding Society* did not exist?
4) What social benefits and impact has *Understanding Society* underpinned, and how?
5) What factors limit the potential for *Understanding Society* to achieve greater impact?

Approach

The evaluation adopted a mixed-methods approach (Figure 1), with a range of activities conducted between December 2021 and March 2022. Seventy individuals were consulted about their experiences of *Understanding Society*. One purpose of the evaluation is to explore how impact arises from using *Understanding Society*. 'Impact' describes both direct and indirect outcomes that result from analysing data and how the structure and coverage of the *Understanding Society* dataset contributes to those outcomes.

Figure 1: Mixed-methods approach



Purposive sampling was used to engage a broad range of stakeholders from a variety of disciplines and sectors. This evaluation focused on direct users of the data, rather than secondary users who engage with published findings or outputs. Top-level sampling focused on academic or non-academic status and was further refined by stakeholders' level of engagement with *Understanding Society*'s data (regular users, irregular users and non-users). Representatives from devolved national governments and international stakeholders were also interviewed.

An evaluation framework – devised by adapting ESRC's Data Infrastructure Strategy theory of change model – guided the development of all data collection tools and analysis. The theory of change model lists the key inputs, activities, outcomes and impacts that the study seeks to achieve. The evaluation framework also guided the identification of pathways to

impact from the field work enabled through using *Understanding Society*. The framework is presented in the technical annex that accompanies this report.

How *Understanding Society* is used

Researchers use *Understanding Society* to explore diverse research questions. No other UK or international dataset has the same depth of coverage or range of data as *Understanding Society*. Its data is used as an important foundation for research and innovation, ultimately contributing to impact.

> "The richness of the variables is very well thought through. I think the modules, all of the areas it covers, is very well thought through and I think it can be used by many scientists. A very rich variety of questions can be answered." ***Academic user – Society, governance & security***

Compared with other UK and international datasets, *Understanding Society* has many unique features and strengths. The extensive range of variables provided is particularly valued. This provides richness and versatility to the types of research questions and analysis that users can pursue.

Academics and non-academics use *Understanding Society* differently. Academics frequently use it to explore their long-term research interests, thus increasing the body of knowledge and capacity to address policy problems. In contrast, non-academic users want to address specific policy questions, often within tight timescales.

Use of specific data, modules, collections, and features

Understanding Society includes different datasets, including the mainstage individual and household surveys, COVID-19 survey, youth data files, the Innovation Panel, teaching datasets, linked data (e.g., to NPD data), and health and biomarker data. In the sample consulted for this evaluation, the individual and household main datasets were the most frequently used, followed by the COVID-19 data files. The general health questionnaire (GHQ) wellbeing variables, and economic and employment data were used most by sampled users, which likely reflects this evaluation's post-pandemic context. **Interviewed users appreciate the capacity to link attitudinal variables to sociodemographic factors and life events**, thus adding value to their research analysis.

Understanding Society data are used at taught master's level and above to develop students' skills in data management, cleaning, linkage and analysis. They are also used by those who teach methods to illustrate quantitative techniques using 'real' data.

The Innovation Panel fosters methodological innovation through the provision of free access to a large sample population and expert support. Users are supported to view the Innovation Panel as a testbed for new methodological ideas.

Special and/or secure license datasets are used to conduct detailed analyses and enable more contextualised conclusions (e.g., by linking to small area data).

The structure of the data from the study facilitates a range of analytical techniques. This includes both relatively straight forward cross-sectional analyses and more complex longitudinal quantitative techniques which can help to establish causal relationships.

Value of *Understanding Society*

Reputational factors

Users, particularly frequent users, are overwhelmingly positive about *Understanding Society*. They say the study has a global reputation and see it as a 'go-to' dataset. The perceived value of the data increases with usage. Many interviewees think that life without *Understanding Society* would have a negative impact at a societal level; they argue that without *Understanding Society*, it would be far harder to address wider social problems and improve people's lives.

Data design, coverage and quality

The **robustness** and **quality** of the data means both academic and non-academic users **trust the data**. This supports researchers to generate **credible findings** that are perceived as statistically robust and hence trusted.

Users value the representation of the whole UK population and the large sample size. The probability sampling within *Understanding Society* enables users to correct for sample and non-sample bias so that the sample can be considered representative.

> "I don't think there's anything, anywhere, comparable that has the same sample size, the whole range of income variables that we need." ***Government analyst***

The large sample size allows users to carry out analysis of relatively **small population subgroups** (e.g., ethnic minorities) which often have policy importance but are not often sufficiently represented in other surveys.

These features of *Understanding Society* support methodological innovation, more complex statistical analysis and provide a testbed for innovative methods. For example, the longitudinal design and data on household structure add additional elements that enable researchers to track participants over time. This enables users to carry out analysis at both an individual and household level from a representative sample of the population as well as considering changes over time and increasing the ability to undertake causal inference.

> "[*Understanding Society*] is unique in that it covers so many sub-groups. It's used by specialists from so many different fields. It's having a major impact on multiple fields of research. There are questions we literally could not have answered without it." ***Academic user - Health & human behaviour***

The COVID-19 files are perceived to be a key strength of *Understanding Society* by all users. They provide timely data about the effects of the pandemic and can be linked to the main datasets to enable the trajectory of experiences and changes during the pandemic to be tracked.

Using the linked datasets[1], particularly the National Pupil Database (NPD), helps researchers to address new research questions by offering multidimensionality to analysis. The main datasets are of primary importance as they provide a 'spine' function that enables linkage.

Learning to use *Understanding Society*

Interviewees say the quality of the support materials and training provision is excellent and exceeds provision for comparable datasets. The absence of this documentation would impact users' ability to carry out their research. Users are supported to navigate the different data waves and variables. The online variable search function is viewed as very helpful, supporting users to gauge the feasibility of different methods/analyses.

> "My favourite tool ever is their variable search online where you can just go to their web page and type in a keyword and it will find you questions that match that and tell you which waves they are in."
> ***Non-academic user – Think Tank***

Aspects of *Understanding Society* that sometimes limit its use

Several common factors limit *Understanding Society's* use, but its strengths far outweigh its limitations. While richness in structure and content is a strength, the complexity of the dataset means that it is challenging to new users, requiring a significant initial time investment to realise the full potential of the datasets. **New users particularly struggle with the overarching design, multiple data files, large number of variables, longitudinal structure and multiple weighting variables**. Once users become competent, the benefits they accrue from using the data mean that *Understanding Society* becomes a key data source in their research repertoire.

Outdated variables are a further challenge cited by a minority of users. Variables designed to measure rapidly evolving fields, such as those associated with climate change, transport/travel behaviour and IT/digital technology may not reflect the current context. This may prevent timely advancements in understanding about certain social research topics. Dropped, skipped or modified variables can be problematic for undertaking longitudinal analysis. There is a balance required between updating content to measure evolving issues, which necessarily means that some variables would consequently be dropped.

Accessing secure and special license files is a barrier for some users facing time constraints, such as non-academic policy users and PhD students with approaching deadlines. Such users are sometimes deterred by the perceived bureaucracy of the application process through UKDS. They typically choose to use data from the standard end-user licence

[1] Linked datasets involve different data sources being brought together so that any individual case has different sources combined. For *Understanding Society*, this means adding external data to the information that is collected as part of the annual survey and interview.

instead. International collaborations may also be impeded where those outside the UK cannot access the secure license files[2].

Social benefits of, and impact from, *Understanding Society*

The unique features of *Understanding Society* help academic and non-academic users achieve a range of academic, social, economic and methodological impacts through their research.

Generating academic impact using *Understanding Society*

Some academics leverage *Understanding Society's* design features to develop new theories and create new knowledge and insight. For example, the longitudinal structure supports causal analysis methods, and some interviewees say causal analysis has increased scientific understanding across a range of disciplines which strengthens evidence-based policy design. The analysis from these researchers has shifted paradigms of how social problems are framed and addressed.

> "*Understanding Society* allows researchers to drill down into the relationship [between employment and mental health], to reduce all alternative explanations as far as possible. That I would say is academic impact. It provides academic and policy debates with more rigorous evidence." **Academic user – Work, education & skills**

***Understanding Society* fosters novel research and new methodological approaches, supporting users to advance scientific methodologies**. For example, economists in America are generating a new forecasting model to predict US military outcomes using *Understanding Society* in a novel way to validate and test their model.

New approaches are made possible for users due to the annual **longitudinal structure**, the **rich combination** of **sociodemographic**, **economic and attitudinal variables**, provision of **social** and **biological data**, the **household dimension**, and the **Innovation Panel**.

Engaging with the data can support academics to publish in **highly rated journals**, which is critical to academic career progression. This is achieved through improved robustness of research and confidence in findings, characteristics provided via the study's longitudinal nature, large sample size and richness of variables.

[2] Since the evaluation, the UKDS has established reciprocal access to secure data via Safe Rooms in Germany, France and the Netherlands as part of the IDAN network, which may contribute to overcoming some of these issues for international researchers based in these countries.

Interdisciplinary research is a **key benefit** of engaging with the datasets, leading to a range of academic impacts. The breadth of variables available supports users' engagement in interdisciplinary work and fosters a **consideration of 'big picture'** questions that explore relationships between different lifestyle variables.

Examples of Understanding Society leading to societal impact include:

- Research demonstrating the number of working hours necessary for positive wellbeing has influenced UK companies trialling a 4-day working week
- Research using distance to relatives has helped inform UK Government's COVID-19 social distancing review and lockdown restriction decision-making
- Public health research showed that the importance of the interconnectedness of health behaviours, which has influenced a shift in paradigm for how politicians think of public health
- Age UK have created an Index of Wellbeing in Later Life using data from *Understanding Society* to provide policy makers and the public with information about what is important for older people's wellbeing

Influencing policy

Understanding Society is a single, reliable source of data analysis of which can influence a range of social and economic impacts. Findings from the data **inform policy, increase public awareness, inform decision-making and change how business/organisational practices operate**. Nearly all interviewees think the data is robust and representative. Using *Understanding Society* as the basis for analysis increases the confidence that policy makers place in the findings when compared to other, less comprehensive datasets. Many policy users say greater confidence in the data increases the chance that their findings will alter policy.

Developing researcher capacity

Methodological impact and increased capacity are achieved when users engage with the training materials and resources. The guidance and training support available on the *Understanding Society* website helps users to increase their quantitative skills and confidence, particularly in relation to longitudinal data skills and choosing appropriate analytical techniques.

All frequent users of *Understanding Society* access the documentation and/or training materials to support their use of the data. Use of these resources, and consequent familiarity with the data, in turn, **increases how much value users place in the datasets** and how much it is used. In the absence of the supporting materials, the datasets may have been overlooked for other less complex data. *Understanding Society* data is **frequently used** in PhD research projects, ensuring the **pipeline of quantitative researchers remains strong**.

01. Introduction

In October 2021, the ESRC commissioned CFE Research to undertake an independent evaluation of Understanding Society: The UK Household Longitudinal Survey. This report is the culmination of a documentation review, 57 semi-structured interviews, 4 focus groups and 12 case studies. The report explores how Understanding Society is used by a range of stakeholders, the impact arising from research findings, as well as users' perceptions of its strengths and limitations.

Background and context for the Review

Understanding Society, the UK Household Longitudinal Study, is the largest longitudinal household panel study of its kind and is one of the UK's leading data collections. *Understanding Society* also incorporates participants from the British Household Panel Survey (BHPS), and continuous data is available for some individuals since the early 1990s. Since 2008, the ESRC has committed over £160million to *Understanding Society*, which is the greatest amount of public funding the ESRC has committed to a single research investment.

The *Understanding Society* dataset is based on survey data from 25,000 households across all four UK nations[3]. *Understanding Society* builds on the previous BHPS that has longitudinal data on approximately 8,000 households since 1991. The study covers a wide range of variables across six core domains: education and training; employment; family, household and social connectedness; civic engagement; income, consumption and wealth; and health behaviours. Data is collected in waves, where individual participants within households are interviewed annually. Another feature of the data collection is the boost samples which occur for ethnic minority and immigrant groups as well as Northern Ireland to ensure that the data is representative of the UK population. Users access *Understanding Society* data from the UK Data Service (UKDS).

The key intended users of *Understanding Society* are those in the scientific academic community, as well as those from policy decision-making organisations and departments. The initial business cases for *Understanding Society* list key stakeholders including

[3] Initially in 2009 when the first data collection phase for the main study was launched, over 40,000 households were contacted to participate in *Understanding Society,* including 8,000 that had previously been included in the British Household Panel Survey. Approximately 40,000 households chose to participate. Business case documentation for Waves 13-15 outline that the number of households is now approximately 25,000, indicating that the number of included households has reduced over time.

devolved administrations, local authorities, research charities and foundations, think tanks, survey developers and some commercial knowledge exchange organisations.

The aims of *Understanding Society* are to track changes in the lives of people, provide unique information on social issues, and support research relevant to expanding scientific knowledge and to the formation and evaluation of policy. The 2020-21 financial year marked the 10-year anniversary of the end of the first wave of data collection. This evaluation sits within a wider body of work which assesses the contribution of *Understanding Society*, its stated benefits and objectives in the context of continued public investment. Specifically, *Understanding Society* produces Benefits and Impact reports and Case Study reports that summarise the impacts arising from work using the dataset. This evaluation supports these documents by incorporating independently gathered user and stakeholder views and perspectives.

Aims and objectives for the research

As one of the UK's flagship datasets, *Understanding Society* has the capacity to help address emerging and pertinent social problems. The ESRC commissioned CFE Research to undertake an independent evaluation of the use, perceptions and impact of *Understanding Society* to inform future business cases and funding decisions. The research assesses how *Understanding Society* is used (and by whom), explores stakeholders' views on the data's perceived strengths and limitations, and identifies the impacts and potential barriers to achieving impact from research using the study's data.

To achieve these aims, the research considers five research questions to gauge the use and impact of *Understanding Society*:

1) **What do stakeholders think of *Understanding Society*?**
2) **How is *Understanding Society* used by stakeholders?**
3) **What have stakeholders been able to achieve which would have been difficult or impossible if *Understanding Society* did not exist?**
4) **What social benefits and impact has *Understanding Society* underpinned, and how?**
5) **What factors limit the potential for *Understanding Society* to achieve greater impact?**

Approach

The research questions are addressed through a rigorous mixed-methods approach that combines an analysis of key documentation pertaining to the benefits and impacts of *Understanding Society* with primary data collection from semi-structured interviews, focus groups and case studies. The participants in this evaluation were direct users (and non- or lapsed users) of the data. This means that the evaluation was interested in

exploring the views of people who are directly using the data, rather than those who use findings or outputs using *Understanding Society*. The exception to this was when speaking to some policy and third sector organisations who had less capacity for primary research and supplemented their own data analysis with other published outputs.

The structure of the report

A thematic structure is used for the report, covering two broad areas – usage and key features of the dataset, followed by an analysis of the impacts achieved. After the methodological summary in Chapter 2, Chapters 3 and 4 cover different aspects of how *Understanding Society*'s data is used by the range of stakeholders consulted. Chapter 3 considers the application of data to research problems, while Chapter 4 discusses what makes *Understanding Society* unique compared with other data that researchers could use, emphasising the methodological impact that has been enabled as a result of the unique features.

Chapters 5 and 6 consider what impacts can arise from research based on *Understanding Society*'s data. Chapter 5 explores academic impacts such as developing academic theories and contributing to the pool of knowledge through producing academic journal articles. Chapter 6 considers the social and economic impacts which arise from applying analysis and findings to policy and societal issues. Chapter 7 outlines the areas that users have expressed can limit usage or impact from *Understanding Society*. These are framed in the context of the broad consensus of support for the dataset and what it can enable users to do. Conclusions are drawn in Chapter 8.

Appendix A contains the 12 case studies that are drawn on in the report and document the impact and benefits of *Understanding Society*. Within the report, boxed vignettes are used to summarise these case studies, with accompanying links. Twelve completed case studies are provided in Appendix A, and boxed vignettes are included in the text where appropriate. Where it is appropriate to document the pathway to impact that was achieved through using *Understanding Society*, diagrammatical representations of the extracted pathway from the Evaluation Framework are included in the text.

The technical appendix that accompanies this document contains the evaluation framework (Annex A) that supports the detailed methodology (Annex B). The discussion guides used for interviews and focus groups are included in Annex sections C and D, respectively.

02. Methodological approach

This section summarises the mixed-method evaluation approach using a documentation review, stakeholder interviews, focus groups and case studies. Appendix B describes the methodology in detail.

Evaluation Framework

An evaluation framework guided the development of all data collection tools and analysis. The framework (Technical Annex A) includes a theory of change model for *Understanding Society* that lists and relates the key inputs, activities, outcomes and impacts that the study seeks to achieve. The evaluation framework was devised by adapting ESRC's existing Data Infrastructure Strategy theory of change model which covers all related ESRC activities. The new *Understanding Society* model was refined in partnership with ESRC during an inception workshop held in November 2021.

Approach

The mixed-methods approach combines a documentation review, 57 stakeholder interviews, 12 case studies and four focus groups. These activities were conducted over a four-month period between December 2021 and March 2022. Seventy individuals across all data collection points (including individual and small group interviews, focus groups and case studies) were consulted about their experiences of *Understanding Society*.

Documentation review

The ESRC provided 28 documents for review. These documents included business cases, progress reports, benefit realisation reports and strategic reviews. All documents were reviewed by the CFE research team, then analysed to assess the type of impact stated (academic/policy/private/public sector) and the beneficiaries of any impact. The analysis was mapped onto the five key research questions.

Stakeholder interviewing

The research used a purposive sample to explore views from across the spectrum of stakeholders. The selection of participants was focused on their academic or non-academic status, and further refined by their type of usage (regular, non-regular, non-user). As part of the sampling strategy, representation was sought from devolved nation governments and international stakeholders[4]. The final sample therefore balanced the different types of

[4] Insights from devolved nations were gained from Scotland and Wales only.

users and disciplines that use *Understanding Society* (Table 1). The constitution of the final sample was partly informed by response rate of users in certain quota groups to being involved in the research.

Table 1: Distribution of interviews by key sample strata

	Regular Users	Irregular Users	Non-Users	TOTAL
ACADEMIC STAKEHOLDERS				
Health & Human Behaviour	5	2	2	9
Society, Governance & Security	5	2	2	9
Work, Education, Skills & Training	4	1	1	6
Environment, Transport & Climate Change	2	0	0	2
Methodological & Statistics	3	1	1	5
TOTAL (Academic)	19	6	6	31
NON-ACADEMIC STAKEHOLDERS				
Government & Public Bodies	8	4	1	13
Private/Commercial, Third Sector & Think Tanks	8	4	1	13
TOTAL (Non-academic)	16	8	2	26
OVERALL TOTAL	34	14	8	57

Information about participants was obtained during the recruitment phase to manage recruitment to each sample stratum. Respondents were asked to complete a short online form which was used to match participants to the sampling quotas. Bespoke interview and focus group topic guides were created for each type of user, and these can be found in the Technical Annex sections C & D.

Case Studies

Twelve case studies have been developed to illustrate the ways in which *Understanding Society* can have impact for users. The case studies were identified through the stakeholder interviews. They showcase the key benefits and impacts in a way that guides the reader through the narrative of how *Understanding Society* is used and how it subsequently translates into impact or outcomes.

Focus Groups

A series of four focus groups were held in March 2022. Each considered a specific concept and topic that emerged as important from the interviews but where more information was needed. These topics were:

- Linking *Understanding Society* to other datasets
- Life without *Understanding Society*

- The Innovation Panel
- Advancing Methodological Innovation

The purpose of the focus groups was to build on the findings emerging from the stakeholder interviews and to assist in building strong narratives around impact to inform case studies and the final report.

Focus group participants were recruited through recontact after an interview and through other free-find methods, including assistance from the University of Essex for Innovation Panel users. A total of 19 participants attended the four focus groups: 12 were academics across a range of disciplines and 7 were non-academic users of *Understanding Society*.

Methodological and measurement challenges

The process of conducting this evaluation elicited some observations that may be useful for future research activities. Following the initial request for participants via an email sent to all users who had downloaded *Understanding Society* from UKDS, a vast number of responses were received. Working with this initial population of users prioritises UK academics and excludes those who access *Understanding Society* through other means and individuals who have chosen not to use the dataset. The lack of complete data on the full population of *Understanding Society* users complicates recruitment and creates possible bias in the sample, and hence this report's findings may not be fully representative of all users' opinions.

Elements of *Understanding Society*, such as the Innovation Panel and the Policy and Partnerships Unit (PPU), are designed to support a sub-set of users. The innovation panel can trial only a limited number of data collection and research innovations each year, so the audience of this feature will be a small subset of main data users. The PPU engages with both data users and evidence users to promote the translation of research into policy and impact. This means there are fewer users in the overall sample that are familiar with these features. For research specifically targeting users of these features, an alternative sampling strategy may be advisable that reflects the niche type of user and smaller overall sample.

A key objective of this evaluation was to explore pathways to impact and to evidence impact that research using *Understanding Society* has led to. Throughout the interviews, users of the data found it challenging to define and operationalise impact. This meant that interviewees often did not think impact had been achieved, despite further probing eliciting some strong examples. Methodologically, competing definitions of impact (such as those used by UKRI councils, in the REF assessment exercise, and within academic and non-academic debate) blur both the understanding of impact and how it can be evidenced. Future research may wish to explore in advance the definitions of impact and how these definitions can translate into terminology and understanding in other sectors. Creating a reliable metric for impact would help ensure that the full dimension of impact is evidenced.

03. Using *Understanding Society*

This chapter summarises who uses Understanding Society and how usage varies between different stakeholders.

Summary:

- *Understanding Society* is used to explore research areas across a wide variety of topics, including health, education, housing, economics and travel.
- *Understanding Society* compares favourably to other UK-based and international datasets – in particular, users value the longitudinal aspect, number of sociological variables and the extensive documentation available to aid users
- The most frequently used datasets are the main individual/household datasets and the COVID-19 data files. Less frequently used are the teaching and youth files. Users can access special and/or secure license files which provides more detail for analysis
- The use of *Understanding Society* data evolves over time as users become more familiar with its structure and content – more experienced users exploit more fully the longitudinal nature of the data, whereas new users may start with only one (or a couple) of the waves in a cross-sectional approach
- As users become more experienced and proficient with using the dataset, *Understanding Society* becomes a key go-to dataset for their research. The range and complexity of analytical methods increases as users become more familiar with the dataset

This chapter explores the characteristics of those who use *Understanding Society* and the purpose for which they use its data. More exploration of why *Understanding Society* is used then follows and the chapter concludes by considering the impact for users if *Understanding Society* did not exist.

Nearly all stakeholders think *Understanding Society* makes an important contribution to research and innovation, enabling data users to answer their key research questions. *Understanding Society* is also viewed as a valuable element of the wider UK data infrastructure that includes other longitudinal studies (e.g., the Millennium Cohort Study and the British Cohort Study) as well as official statistics (e.g., Office for National Statistics datasets). Stakeholders are overwhelmingly positive about *Understanding Society,* regardless of how much they use the data or their user role (academic and non-academics).

Who Uses *Understanding Society*?

Overview

The documentation review outlines that most UK-based *Understanding Society* users (>90%) are from UK higher education institutions[5]. The remainder are UK analysts from the Government (4%), the third sector (1%) and the commercial sector (1%). Many different academic disciplines use *Understanding Society* data in their research. The following table summarises the variety of research areas and topics that interviewees used *Understanding Society* to address.

Table 2: Specific topics that interviewees outlined as using *Understanding Society* for

Participants' overarching Research Area (Quota)	Specific sub-topic
Society, governance and security	Poverty, housing, transport, economic security, political inclusion and engagement, migration, social integration, ethnicity, social mobility, social security, family dynamics, household finance, mortgage debts, bereavement, electoral behaviour, cultural consumption.
Health and human behaviour	Testosterone and behaviour, epidemiology, mental health, health statistics, ageing, carers and carers' experiences, stress and allostatic load, healthcare, loneliness, public health, multiple risk behaviours, physical environment and health, life satisfaction, smoking behaviour, family complexity and children's health, alcohol and drug use
Work, education, skills and technology	Special educational needs and disability provision in schools, educational outcomes, loss of earnings due to COVID-19, welfare and work, flexible working practices, job satisfaction, labour economics, unemployment, workless families, career guidance, insecure working conditions, remote working, unpaid work and inequalities,
Environment, climate change, energy, transport and urban	Urban inequalities, spatial segregation, neighbourhood effects, migrant integration, global travel analysis.
Methodological research, statistics, data science	Statistics, promoting quantitative methods, large social surveys, big data, modelling.

Comparison with Other Datasets

***Understanding Society* has advantages over other UK-based datasets.** Interviewees commonly compare *Understanding Society* to ONS datasets (the Labour Force Survey and the Annual Population Survey), other long-term datasets (British Election Survey, English Longitudinal Survey of Ageing (ELSA)), and surveys conducted by commercial research agencies such as YouGov.

[5] These figures are calculated using data download figures collected by the UK Data service.

Compared with some ONS datasets (e.g., the Labour Force Survey (LFS)), in any given wave *Understanding Society* has smaller cross-sectional sample sizes, but the range of variables across various dimensions of people's lives makes the dataset richer and more versatile. The tracking of people over time is possible with longitudinal data in *Understanding Society*, whereas the largely cross-sectional nature of LFS means that the same individual cannot be followed through their life course to see how they respond to particular events/occurrences[6].

Compared with other longitudinal datasets, *Understanding Society*'s larger sample sizes and data on household structure add an important element to tracking change over time.

***Understanding Society* compares favourably with international comparator datasets**, and its reputation is global. Internationally, *Understanding Society* is compared with the German Socio-economic Panel (SOEP), the Household, Income and Labour Dynamics in Australia (HILDA) survey and the Panel Study of Income Dynamics (PSID) (amongst others). Interviewees emphasised the depth of variables and the quality of documentation exceeds these comparison datasets.

> *I believe that Understanding Society is the envy of the world. When I have presented at international conferences, colleagues have come up and said, 'Wow, we wish we would have something like that in our country,' and I say, 'Well, data go back to 1991.' They can't quite believe that it's true, but it is.*
>
> **Academic user – Work, education, skills & technology**

Different Experiences Depending on Sector and Capacity

***Understanding Society* is used differently by academics and non-academics**. Academic users explore their research interests using *Understanding Society* data, often during projects lasting many months or years. Using the data over a prolonged period helps academics incrementally develop their expertise and competence with the dataset.

Conversely, non-academic users seek answers to specific policy questions, often to tight timescales. Consequently, non-academics may have less time available to learn how to use the data, although those with past academic careers apply prior learning. For answering immediate policy questions, the time required to analyse *Understanding Society* can reduce the likelihood that it will be the go-to dataset.

> *I don't think we could quickly respond to emerging immediate policy questions anyway [using Understanding Society] because it's like, 'can you get me an answer in a day?',*

[6] The Labour Force Survey does have a longitudinal element but it is only possible to track the same individual for five quarters.

> *rather than, 'you can spend a couple of weeks cutting the data and linking it, and then find the answer to the question.*
>
> **Non-academic user – UK or devolved government**

How do stakeholders use *Understanding Society*?

Datasets used

Interviewees use the individual and household main datasets most frequently, followed by the COVID-19 data files. A further breakdown of the use of different datasets can be found in Table 2. The wellbeing GHQ variables and economic and employment data during COVID (e.g., job retention/furloughing schemes) were most commonly used by interviewees. These variables can be linked to other sociodemographic factors and life events, which adds value to *Understanding Society* over more specific topic-based surveys which may have more detailed variables but lack the supporting demographic information.

Table 3: Number of interviewees who have used different *Understanding Society* datasets

Standard Individual and household files	37	Teaching dataset	5
Youth Data files	7	Secure License files	6
Nurse & Biomarker data	6	Special License files	18
Innovation panel data	5	COVID-19 Survey dataset	23

Note: The data in this table is the number of users who have downloaded and used the data for research and/or teaching.

The teaching and youth data files are used least by interviewees. Teaching using *Understanding Society* usually occurs at taught Masters level and above. Those who teach at this level emphasised that encouraging students to work with the data themselves can develop the skills needed at that level (e.g., cleaning data, linking waves together). Undergraduate programmes are less likely to teach these skills. Those who teach methods and methodology often use *Understanding Society* datasets and analysis to illustrate quantitative techniques and methods using 'real' data:

> *I usually use Understanding Society to facilitate students' interest in how relevant interesting topical quantitative data can be. Understanding Society encourages students to see that quantitative data are important, and what you can do if data like this is available on tap.*
>
> **Academic user – Work, education, skills & technology**

For those conducting research about children and young people, there are some concerns about missing data and poor response rate in the youth files. Moreover, interviewees emphasise that specifically for young people, there are many alternative datasets (including the Millennium Cohort Study, Next Steps etc.) that have longitudinal birth cohort data which may be more useful for research questions specifically relating to the experiences of young people in isolation. However, strengths of the youth data files include the ability to link to wider household information and structure.

The COVID-19 files are seen as a key strength of *Understanding Society*. They provide timely data about the effect of the pandemic from a large sample during a critical time. Researchers appreciate the link between the COVID-19 and the main datasets which allow analyses of COVID data variables by lifestyle and demographic factors collected at earlier waves. This means that researchers can draw on a history of previous information and data about cases to explore the trajectory of their experiences during COVID-19. Users comment that the COVID-19 files are unique and provide rich data in a timely fashion. An example of impact from the COVID-19 dataset can be viewed in the vignette below. As one user summarises their thoughts about the COVID-19 files:

> *I thought the COVID survey was spectacular – to have those regular surveys going through COVID and to be able to link those people to the prior COVID period, I think is something that almost no other country has. And what we can see there is all the effects of COVID, the lockdown.*
>
> **Academic user - Society, governance & security**

CASE STUDY 1: Using *Understanding Society* to provide evidence on family and household behaviours to inform Government COVID-19 policies

The *Understanding Society* COVID-19 dataset has been used extensively to inform government policy and decision making in response to the pandemic. For example, The Office for National Statistics (ONS) used data on household relationships, travel times and living relationships to inform the government's decision making on social distancing in 2020. Analysis of the COVID-19 survey data informed a review of the social impact of COVID-19 on ethnic groups, and further data from the COVID-19 survey was one of several sources submitted to the Government as part of the social distancing review. The timeliness of the COVID-19 data, coupled with the ability to track changes in individual and household behaviours over time, through linking to the main dataset, enabled the ONS to use *Understanding Society* to inform these policy decisions. These unique features also allowed the ONS to predict which behaviours may change if the government implemented (and adapted) social restrictions.

The diagram below indicates the pathway to impact that has been enabled through using *Understanding Society* for this methodological impact purpose, as applied from the Evaluation Framework (Appendix A). Similar pathways to impact are presented elsewhere in the report.

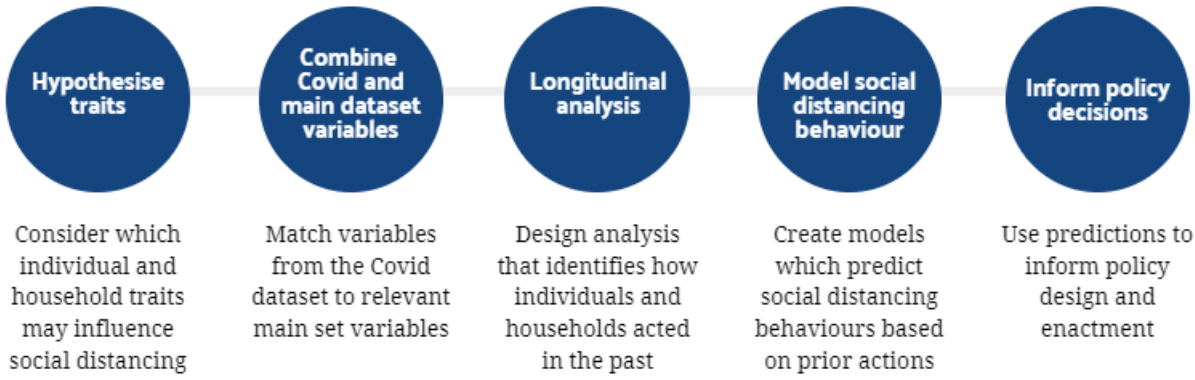


Figure 2: Using *Understanding Society* data to inform COVID-19 policy

Using the greater detail and information provided in the special and/or secure license datasets means that interviewees can conduct analyses with greater granularity and make more contextualised conclusions. Users noted that accessing the special license files (for lower-layer super output areas (LSOA) data) and secure license files (for grid reference) allowed them to incorporate greater levels of detail into their analysis. Despite the strengths in the resulting analysis from extra geographical details, some users acknowledged that the time to access special license data can be problematic for quick research outputs.

The facility to link with other datasets is perceived as both a strength and limitation of *Understanding Society*. A key linkage valued by interviewees was with the National Pupil Database (NPD). For educational researchers and those interested in young people, the potential of linking to NPD data is a strength. However, most of the researchers who use this data have found that the sample sizes of matched cases consenting to linkage are too small for meaningful analysis. Further analysis of users' perceptions of the linked datasets can be found in subsequent chapters.

Users highly value *Understanding Society* for allowing them to produce robust and credible findings. The extensive structure, including multiple waves, large numbers of variables and large sample size, means that the time to become familiar with the dataset can be lengthy. To enable their employees to use the data quickly and accurately, some organisations have created their own extraction software to ensure the data can be used efficiently to answer pertinent research questions. Organisations have developed innovative methodologies to ensure capacity building for their employees as well as ensuring data efficiencies for conducting research and evaluation. This is explored more in the case study vignette below.

CASE STUDY 2: Using *Understanding Society* to build quantitative skills and increase productivity

Non-academic organisations highly value the data within *Understanding Society*, as it enables them to answer important and often complex research questions. The large sample sizes, wave structure, and range of variables in *Understanding Society* requires considerable knowledge and skills to use it proficiently. Because of the high value placed on the data in *Understanding Society*, the Institute for Fiscal Studies (IFS) have created their own extraction software to build internal analysis capacity and the quantitative skills base to support research about household income and spending behaviours. The data extractor helps increase quantitative capacity through enabling users to merge files, recode family member relationships, create consistent definitions of variables of interest, and merge cross-wave information that promotes longitudinal analysis. Greater confidence in using the data helps to increase the credibility of findings for the IFS and means that outputs remain at the forefront of advancements in knowledge.

Analysis conducted

Interviewees showed that the main immediate value of *Understanding Society* comes from users directly interrogating the data. This means that users conduct analyses on the dataset rather than applying findings from existing analyses of *Understanding Society* to their research questions. The nature of those included in this evaluation, as users of the data, meant that all of those interviewed are interested in primary analyses of the data so that they can match it to their research questions and apply their perspective to the data. The extent and richness of variables means that different combinations of variables or analyses can be conducted for a given topic. Different users (both academic and non-academic) have their own requirement to include certain variables or demographic groups (age/gender/ethnicity) that can be best addressed by conducting their own analysis.

Quantitative analytical skills evolve and are strengthened for long-term users. Interviewees acknowledge fewer advanced quantitative skills were needed at the start of *Understanding Society* when there were fewer waves and analysis required less 'stitching' together of data:

> *The first project I did with [Understanding Society] I treated it more like a cross-sectional, but now I use it mainly as a panel, making use of the longitudinal aspect.*
>
> **Academic user – Work, education, skills & technology**

The large number of datasets are intimidating for first-time or inexperienced users and can limit their confidence to complete longitudinal analysis or more complex modelling.

There is some evidence that smaller private and third-sector organisations may commission others to complete primary research using *Understanding Society*, and then use their findings within their reports. Smaller organisations commission longitudinal analysis because in-house capacity for such work is limited.

Stakeholders differ in their views about the 'best' analysis to achieve impact with their work. Some stakeholders emphasise the importance of longitudinal analysis and the impact of key events on subsequent behaviours and situations. In contrast, other users suggest that funders and policy makers may be less familiar with longitudinal analysis and perceive cross-sectional findings to be compelling and insightful about the current issue:

> *One of the reasons I think cross-sectional data analysis is so popular with policy makers is that's entry level, and doesn't require as much analytical knowledge.*
>
> **Academic user (focus group – Health & human behaviour**

***Understanding Society* enables complex analyses,** such as causal analysis which can more convincingly be achieved through longitudinal analysis. Interviewees noted survey waves before and after key social events (such as Brexit, COVID-19, bereavement, change of life circumstances) facilitates better analysis of causality. This is because the influence of key events can be tracked, and potential causes of changes to other variables before and after can be explored. The extent to which more complex analyses are possible is provided by the following academic user and covered further in Chapter 4:

> *The longitudinal data…allows you to have more sophisticated research designs than a simple cross-sectional survey…you can be more confident that your findings might be closer to being causal, rather than just associations.*
>
> **Academic user – Work, education, skills & technology**

As *Understanding Society* has matured, and more waves of data are available, more complex analyses are possible. With more data, users are now able to "realise the full potential of it" (academic user – Work, education, skills & technology) and apply more advanced quantitative methods to explore the variables and their relationships within the dataset.

Discussions during one of the focus groups (Life without *Understanding Society*), centred on the use of the dataset to self-validate measures of health, adding robustness to analysis. The multiple variables and alternative measurements within the same study increases the robustness and confidence that researchers have in their findings:

> *When you talk about health, you can measure it using different validated measures and they're in the same dataset. So not comparing to other datasets, but within itself. Wellbeing, for example, using different measures within Understanding Society and they're all validated. It provides you with alternative measurements, which is robust, and multiple robust measures in the same dataset.*
>
> — **Academic user (focus group) – Health & human behaviour**

Innovation Panel

The Innovation Panel is a sample of 1,500 households that is used by researchers as a place to test innovative ways of collecting data and develop new areas of research. Users apply in an annual application process to have their study included in subsequent waves of the panel. Across the main interviews, only two users had applied to have a study in the Innovation Panel, with the rest of the interviewees having little knowledge of the panel, its purpose or how it works. The Innovation Panel is used by a niche subset of users (often with a methodological focus) and those working on the design and delivery of the study for *Understanding Society* due to its unique purpose. During the Innovation Panel focus group, one user expressed their opinion that the Innovation Panel "is a victim of its own uniqueness [...] it doesn't exist anywhere else, you wouldn't expect to be able to do something like what the Innovation Panel offers" (academic user – Health & human behaviour).

For those who have engaged with the Innovation Panel, the strengths include the ability for researchers to access a large sample population for bespoke survey questions (generally focused on testing methods questions) that is free and cost-effective to researchers.

> *If I'd had to do it from scratch, it would have cost a ridiculous amount of time and money, but also, it was one question I wanted to ask, and to set up a whole study to just put that one question in, you wouldn't set up a whole study just to do that.*
>
> **Academic user – Health & human behaviour**

CASE STUDY 3: Using the Innovation Panel to foster methodological innovation

A team of researchers applied to have their study included in the Innovation Panel to test out a novel methodology to define and measure successful ageing (the criteria determining whether people are living well in old age). Research participants, via the Innovation Panel were provided with three vignettes that required them to rank how successful the vignette characters were at ageing. Using the vignette method with a large sample enabled a novel data tool to be developed, to bridge the qualitative-quantitative divide in the study of ageing. As a result of this research, a heavily cited article has been published and the findings have contributed to a shift in how policy and medical professions view successful ageing – emphasising the importance of both the physical and social aspects of ageing.

Users of the Innovation Panel noted that the application process and designing the study was straightforward, supportive and accessible to researchers: "because the support was so good, you don't need to be an experienced researcher. You need to a have a good idea, that's the bottom line" (Academic user [focus group] – Health & human behaviour). For researchers who apply to have a study in the Innovation Panel, there is a perception of a freedom to be creative due to a perceived lack of risk. This is because researchers feel that the success of the Innovation Panel is to explore new methodologies and is not dependent on the study itself being successful in producing 'good' data. This is different to 'normal'

funded research studies that are required to produce outputs and demonstrate impactful findings. Users of the Innovation Panel therefore felt that it offered a safe space to experiment and try new ideas:

> *The vignettes were something we hadn't used before, we didn't know how acceptable it would be. We got really good responses, I think nearly everybody did it. The risk wasn't ours. If it had crashed and burned and only 5% of people answered, that wasn't on us. We didn't apply for funding, so we didn't have to say to anyone we failed. The risk was taken from us and it was successful. The method was innovative, but there was no risk in using it which was really good.*
>
> **Academic user (focus group) – Health & human behaviour**

Policy and Partnerships unit

The Policy and Partnerships Unit (PPU), as part of *Understanding Society*, works with government, charities, think tanks and business to enable the effective translation of research and evidence into impact. Across the interviews, **awareness of the Policy Unit is low**, which may reflect the prioritisation of academic users within the sampling quotas, who are not targeted users of the Policy Unit. A specific evaluation of the PPU would use different sampling techniques. Therefore, whilst awareness of the PPU by typical users is low, this is not necessarily unexpected. Nevertheless, improving the visibility of the PPU (e.g., on the *Understanding Society* website) would be effective in increasing awareness and potential use of the facility throughout different users of *Understanding Society*. Even amongst the non-academic portion of the research sample, awareness of the PPU remained low. Only one non-academic interviewee had engaged with the policy unit in a previous role in a Government department. This individual sought assistance to produce and deliver conference presentations and disseminate research findings. For many third sector or commercial organisations, interviews highlighted that they have their own policy units or routes to impact, suggesting that this may negate their need for using *Understanding Society*'s policy unit. These users were therefore unclear what the PPU could add to their work.

04. Unique features of *Understanding Society*

Understanding Society has unique features that attract users to the dataset. These features enable methodological innovation and impact for users. For experienced users, Understanding Society can enhance their methodological capacity and quantitative skills. This chapter explores these unique features and the methodological impact that can arise.

Summary:

- *Understanding Society* has unique features that have established it as a leading dataset in the UK
- The longitudinal design, large sample size (including ethnic boost samples), large number of variables and household structure are the key elements that users consider are strengths in supporting complex analyses, allowing researchers to address key social problems from an interdisciplinary perspective
- The large number of variable topics and design features of the dataset allow for innovation in methodologies and analysis. The dataset can be used in teaching to advance understanding of methodological perspectives and bring analytical methods to life
- The documentation provided for *Understanding Society* is comprehensive and enables researchers to learn to use the dataset, advances their quantitative capacity and encourages more advanced analyses
- Interviewees think the quality of research conducted would be significantly weaker in the absence of *Understanding Society*, which can have implications for policy

Across the interviews, all interviewees expressed extreme appreciation and praise for *Understanding Society*. The strengths of the dataset were frequently mentioned and significantly outweighed any weaknesses – both in number, and in intensity of feeling:

> *Understanding Society is a great national treasure. We should celebrate it and not be shy about it, boast about it, tell the world, and let the world use it.*
>
> **Academic user (focus group)– Society, governance & security**

This chapter explores the perceived strengths of *Understanding Society*. Chapter 7 presents perceived challenges with achieving impact as a result of some elements of *Understanding Society*. Some features of the dataset are viewed as both strengths and weaknesses. In these instances, both perspectives are presented, whilst also indicating the relative weighting allocated to each viewpoint. A visual representation of the features that are both strengths and weaknesses can be viewed in Figure 3. It is important to note that in this figure, the relative strengths are measured against the user group – for instance, where 'most' academic users have stated an element as a positive, but only a 'few' non-users have stated it as a limitation, this can be evaluated as a predominantly positive aspect. The figure shows that nearly all the areas where there are perceived weaknesses are primarily problematic for new and non-users; where the enablers become realised once an individual is experienced and comfortable with the dataset.

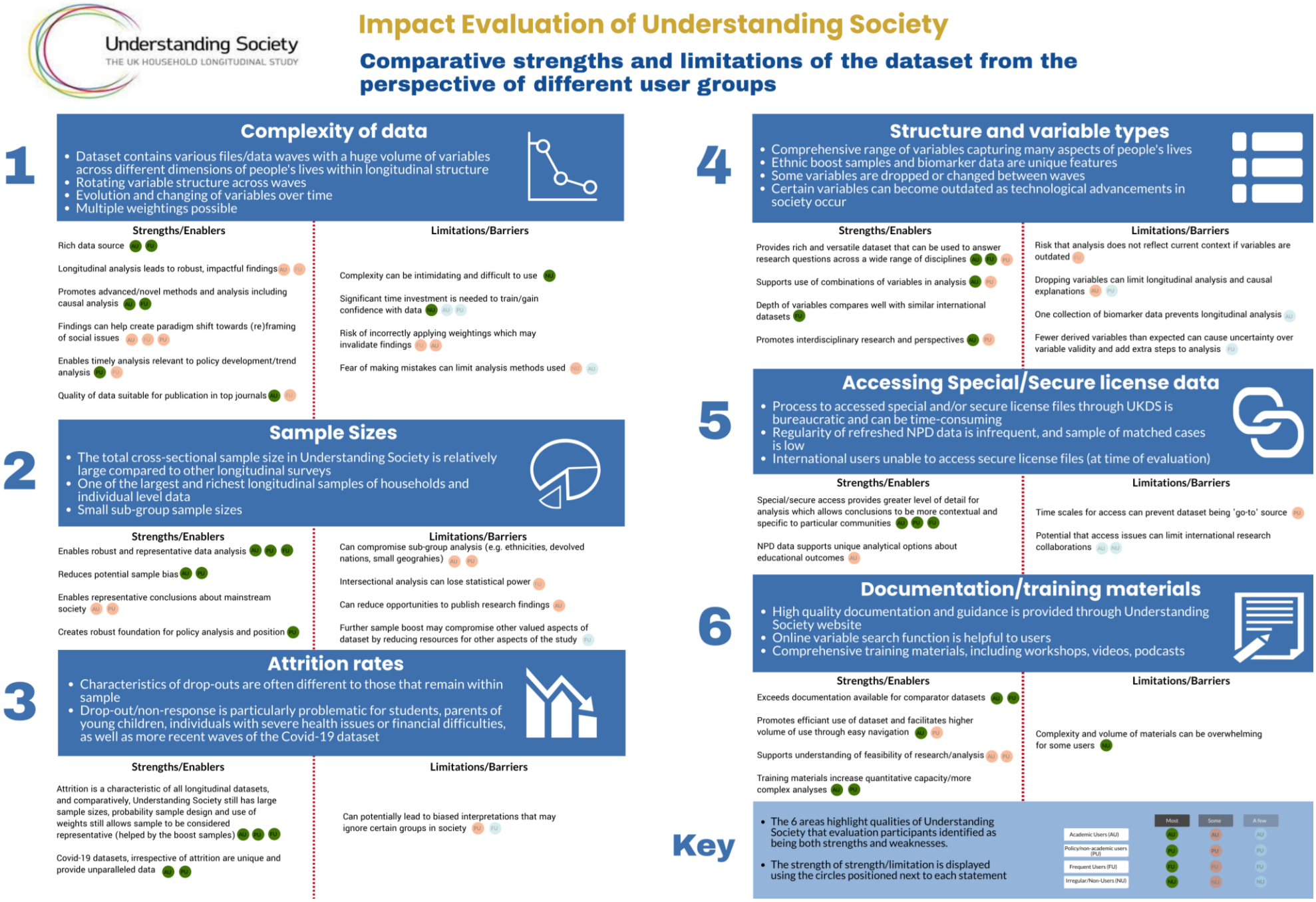


Figure 3: Infographic of relative strengths and weaknesses of *Understanding Society*

Why do stakeholders use *Understanding Society*?

Robustness and data quality

Nearly all stakeholders think *Understanding Society* is a valuable, essential and high-quality dataset. Many stakeholders note ***Understanding Society*'s data collection methodology is statistically robust. Users therefore trust the data** and are confident in their findings. Many respondents said their research/project was viable in the absence of *Understanding Society*. However, three in four of this group say the quality of their analysis would weaken because less robust methodologies are used to collect data for alternatives.

> *I would have been able to do the research, but it would not have been of the same quality. I would not have been able to trust the results of my analysis as much as I am able to trust Understanding Society data.*
>
> **Academic irregular user – Methodological research**

The **national and international reputation** of *Understanding Society* is important for users to feel confident about the conclusions they draw. The "brand name" (non-academic user – Independent researcher) is powerful in encouraging more researchers to use the dataset. This is because users trust the dataset and are confident in the conclusions that are possible. Analysis using *Understanding Society* is credible in policy and academic discussions.

Many academics say *Understanding Society* is an invaluable resource and a comprehensive dataset which addresses a range of research questions and interests. Comprehensive describes several attributes. The whole UK population is represented in the *Understanding Society* data collection design. This allows researchers to conduct accurate analysis of their research questions for the population, or specific sub-groups. The large overall sample size results in usable sample sizes for many sub-groups. The ethnic population boost sample is mentioned by several interviewees as valuable for targeted analyses, as are analyses by age, gender and socio-economic background.

> *The fact that you have [an ethnic] boost sample means that you have more cases and you can do more things with the data and basically, in statistical models, it means you can also use more variables and you can disaggregate more because you have this good sample.*
>
> **Academic user – Methodological research**

However, a small number of users specifically looking at the field of ethnicities (often with intersectional analysis against other social factors) feel that a larger ethnic boost sample would improve sub-group analyses for individual ethnicities where the sample is too small for meaningful analyses.

***Understanding Society* covers many concepts with underlying demographic variables which enables interdisciplinary research.** For instance, users interested in travel describe analysis using variables relating to households' employment, and sociodemographic and financial characteristics. Such a combination of variables is not present in other travel-related datasets, as highlighted in the vignette below.

CASE STUDY 4: Adopting an interdisciplinary perspective using *Understanding Society* to meet the shared needs of research and policy

The unique features of *Understanding Society* enable researchers to address topics from different interdisciplinary perspectives and simultaneously meet the shared needs of research and policy. For example, recent travel and transport research, carried out by an academic team and Government department used similar variables to approach the same topic from different angles and with different purposes. The Government department used *Understanding Society* findings to show the value of travel investment on people's lives, which is appealing to policy makers. Academic impact has also been achieved – for example one paper focused on the tension between migrant travel behaviour and climate change. A further output advanced scientific theory by exploring potential causal mechanisms underlying higher levels of long-distance travel activity among residents of urban areas.

The wide range of variables in *Understanding Society* enables researchers to look at social problems from a variety of perspectives. Interdisciplinarity is enabled because users can combine novel variables on a range of different topics with underlying social demographics to explore links that may not have been considered before:

> *Understanding Society gave me an opportunity to link air travel behaviours to other aspects such as migration background, and social network dispersion, so whether your friends and family live abroad, where your friends live. The link between those two things, it turned out no one had explored that… by having these different modules, you end up having different questions and variables together that you normally wouldn't have and that makes for a lot more possibilities for analysis than otherwise.*
>
> **Academic user – Environment, transport & urban**

Interdisciplinary knowledge is crucial for researchers to be able to make the most of the unique features of *Understanding Society* for innovative research. For example, researchers working on case study 5 (see below) noted that they were able to exploit the valuable features of *Understanding Society* by using 'domain experts' to help the team understand elements of the underpinning topic and variable definitions, meaning that their "algorithm has a head start" in being able to successfully predict outcomes (Non-academic user – International research institute).

Demographic variables allow researchers to control for relevant confounding factors in regressions. These findings can then be used to make specific policy and practice recommendations or consider 'big picture' questions that explore relationships in different

lifestyle factors. Many users take advantage of this to design cross-disciplinary research or apply analytical methods that identify relationships between disparate concepts.

> *[Understanding Society] is unique in that it covers so many sub-groups. It's used by specialists from so many different fields. It's having a major impact on multiple different fields of research. There are some things we literally could not have answered without Understanding Society.*
>
> **Academic user (focus group) – Health & human behaviour**

To this extent, the range of variables in *Understanding Society* has been instrumental in enabling one international research team to use the dataset to forecast life outcomes, as a means of validating their own forecasting model, which operates in an entirely different context (of military resource planning). As a methodological impact, this example demonstrates the use of the dataset for innovative methods and for challenging current methodological approaches. These researchers used the variables in *Understanding Society* to identify potential confounding variables and control for them in statistical modelling:

CASE STUDY 5: Validating and testing a forecasting framework using *Understanding Society*

An example of methodological impact is showcased via work carried out by the Institute of Defense Analyses in Washington DC, who have developed their own forecasting models (the Finite-Interval Forecasting Engine – FIFE) to inform military force planning. All waves of *Understanding Society* and the BHPS were used as an external dataset to test and validate the FIFE model. Researchers applied a machine learning algorithm to forecast outcomes up to four waves ahead. Testing proved that the model was effective and produced high performing individual-level forecasts. The Institute of Defense Analyses are now using the knowledge gained from the validation of this model to train new models to inform a controlled trial of service member retention initiatives. This demonstrates the contribution of *Understanding Society* to develop innovative advancements in statistical methodologies.

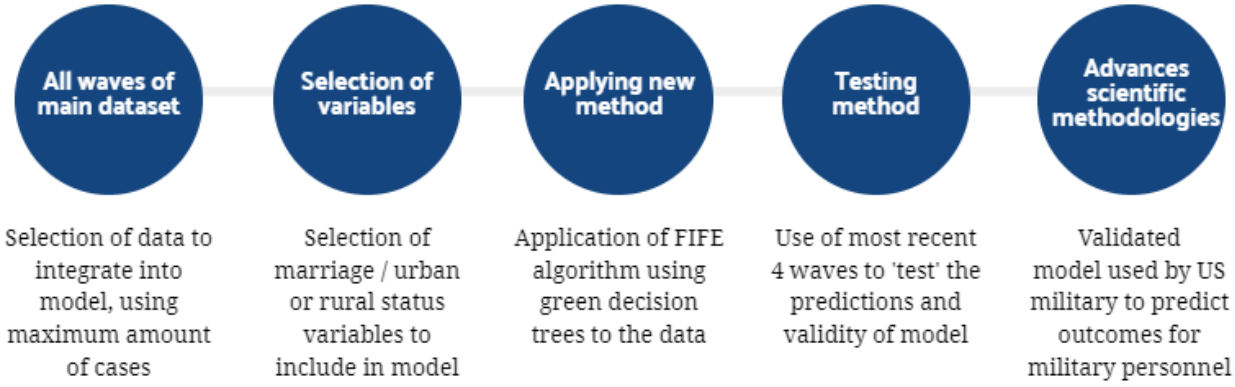


Figure 4: Validating forecasting models

The different variables and rotating modules allow for creativity when answering research questions that may not be possible in other datasets. This is because *Understanding Society* users can explore relationships between variables that are not commonly collected together and reframe questions within different contexts.

> *It has an element almost of serendipity, when you have all those rotating modules, suddenly, you can link that information, even though maybe no one really had in mind to encourage researchers to make that link and then someone comes up and says, 'Oh, I could do something with that.' I think it induces a lot of creativity.*
>
> **Academic user – Environment, transport & urban**

Interviewees note that whilst alternative panel datasets are available, they do not provide the same richness that users value in *Understanding Society*. **Without the rich variables in *Understanding Society*, the significance of the conclusions and findings would be limited, resulting in less impactful research.**

Several users say the sample's size and representativeness means that researchers can make "conclusions about mainstream society" and **contextualise their findings** (non-academic user – independent research institute)**.** The large sample size has implications for how *Understanding Society* is used with **senior stakeholders and policy makers** – it enables them to draw meaningful inferences from data offering better evidence for decision-making.

> *With Understanding Society, one of the things that's quite impactful is purely sample size. It's very big and that's very helpful because it lets you extrapolate to general population more without having to take three averages that's generally what senior stakeholders like.*
>
> **Non-academic user – Charity sector**

Users express a concern that **life without *Understanding Society* would impact at a societal level**. There is a perception that the findings from *Understanding Society* can contribute to addressing wider social problems and help to make improvements to impact on people's lives:

> *I think the UK would be much poorer place if we didn't have [Understanding Society] to be able to generate that evidence base to live a better society, I think for everyone.*
>
> **Academic user - Methodological research**

Interviewees noted that the **structure of *Understanding Society* reduces potential biases in the sample, including non-response rate**. The longitudinal nature means that researchers know more about the people who do not respond, meaning that these characteristics can be factored into analysis, correcting for any bias. Due to being able to

correct for biases, any conclusions or policy recommendations are based on a sample which more accurately reflects the wider UK population, thus increasing the relevance of findings from *Understanding Society*.

Large sample size allows methodological innovation. The presence of larger sample sizes means that more complex statistical analysis can be conducted, as well as providing a testbed for innovative methods. Increased awareness and comfort with using the dataset means that users can access greater types and complexity of analysis methods. Accessing the training materials enables users to be more comfortable in making sense of the data and to "think about using the data in more creative ways" (non-academic user – Charity sector). A particular methodological impact arises from the large sample size because users can use more robust methodologies:

> *I think methods-wise, because you have got a big data set you can test out different methodologies, and we were using cluster analysis. Because we had got all these variables from different data waves, we had something like 15,000 cases, so, we could randomly sample within that and still get a really good robust representative sample. Because you're always afraid when you start cutting from data sets with different variables that you are introducing bias.*
>
> **Academic user – Environment, transport & urban**

Longitudinal structure and household data

The longitudinal design enables researchers to track participants over time and hence how participants' lives change. A few interviewees note that the value in the data accrues as the longitudinal element matures. Combined with the breadth of variables in the dataset, the longitudinal design also creates opportunity for analysis of causality. For instance, the panel structure of *Understanding Society* enables researchers to control for both observed (through the variables) and unobserved individual characteristics that influence the relationship between, for example, mental health and employment. This is important for attempts to establish causality. Household characteristics (composition, income, social characteristics, etc.) further enhance the analytical possibilities for several users. A few users say they developed a greater appreciation of the complexity in which individuals experience their lives.

> *It enables you to actually look at things in these households and household change over time in a way that wouldn't just be focused around one individual, it would actually be about people entering and leaving the household so that's a major aspect of the study.*
>
> **Academic user – Work, education, skills & technology**

The longitudinal nature allows for a range of advanced methodologies and analytical methods to be used. For instance, researchers have used cluster analysis, algorithmic methods, forecasting and regression modelling. The benefit of using more advanced and

innovative methods is that it allows researchers to 'look at where people really are' (academic user - Society, governance & security), understanding the different facets and dimensions of their lives.

Linked datasets

Data in *Understanding Society* can be accessed in formats that has been linked to other data sources, including DVLA, National Pupil Database (NPD), geographical indicators and (for the COVID-19 modules) administrative health records. Access to these datasets is only possible through secure and special licenses. Across the interviews, the majority of users of linked datasets were educational researchers who were most interested in the linked NPD data. For researchers who have used linked datasets, they perceived the benefits as increasing the robustness of their results, providing opportunities to answer new and novel research questions, and providing links between subjective and objective measures.

Data from the focus group specifically on the use of linked datasets emphasised that the potential for linked datasets is significant, but that emphasis must also concentrate on the main dataset primarily:

> *It's good to invest in linked datasets, but we need to keep the focus on the main datasets so that it can function as a kind of spine around which administrative data can be linked in. There aren't many surveys that have that spine function. Remembering the core of Understanding Society is really important, as it's the basis of the quality upon which any linkage can occur.*
>
> **Non-academic user (focus group) – Independent research institute**

Factors that enable users to work with *Understanding Society*

Documentation and training materials

The documentation available allows users to learn how to use the datasets and become familiar with the unique features within *Understanding Society* (see also Chapter 3). Nearly half of the interviewees comment that the **documentation available is excellent** and helps them navigate the different waves and variables. Several interviewees praise **the online variable search** because it helps researchers gauge whether potential research is feasible. The online search helps users design statistical models to interrogate hypotheses. Variable searches include respondent counts that allows researchers to make some analytical decisions before interrogating the data.

> *My favourite tool ever is their variable search online where you can just go to their web page, type in a keyword and it'll find you questions that match that, tell you which waves they are in, it'll give you the exact wording and it'll give you a breakdown of how many respondents picked each option for that question. It tells you who it was asked to, the universal respondents, it's just so cool and I really wish that other surveys had that.*
>
> **Non-academic user – Think Tank**

The good level of documentation means that more research has been conducted using the dataset. For some researchers, the documentation is critical for the existence of projects and research:

> *If the documentation was only half as good, I don't know if I could have done the project.*
>
> **Non-academic user – Independent research institute**

In particular, the information about the variables and how they changed across different survey waves is critical for allowing this research project to take place.

Frequent users of *Understanding Society* are very positive about the extensiveness of the documentation; it is viewed as a complete record of all data available. In contrast, less frequent or inexperienced users are sometimes overwhelmed and confused by the volume of the documentation. Some suggestions about the limitations of the documentation can be found in Chapter 5.

All frequent users of *Understanding Society* access the training materials and documentation to support them in their use of the data. Approximately half of all interviewees used *Understanding Society* training materials in the form of courses, podcasts, seminars and videos. Approximately half of the remaining interviewees say they learned how to use the dataset from reading documentation and other supporting information. The rest do not analyse data themselves or are not users of the data.

Use of the training materials and resources increases users' quantitative skills and their confidence in using *Understanding Society*. Users of the training materials emphasise that they feel more confident in knowing what options they had for analysis, how to link the datasets and how to use statistical software:

> *I did a 2/3 day workshop organised by Understanding Society specifically for people in third sector organisations. That was super useful, and it was only two days but my skills and confidence with the dataset was so much improved…before I muddled my way through, but [after the training] I felt so much more confident and more aware of the wider range of things you can do with it.*
>
> **Non-academic user – Charity sector**

An important impact arising from the use of resources and documentation is the increased use of the dataset. As users access the training and become more confident with their dataset, they increasingly view *Understanding Society* as a valuable resource. Users develop knowledge of what is possible within the dataset, increasing its perceived value to them. **Without the training and documentation, *Understanding Society* may have been overlooked for other, less complex datasets.**

> *'I definitely think that if I'd been faced without all the resources that are available, I don't think I would've used it at all, and I don't think it would've been a factor in terms of us thinking about it as a resource at all'.*
>
> **Non-academic user – think tank**

05. Achieving Academic Impact

Understanding Society is primarily used by academics in UK higher education institutions. This chapter explores the extent to which the datasets support achieving academic impact and the mechanisms that facilitate this.

Summary:

- Through using *Understanding Society*, academics have developed new theories and used causal analysis methods which have led to shifts in scientific understanding within disciplines
- The longitudinal design of *Understanding Society* allows causal explanations to be developed
- The innovation panel, wave structure and combination of sociodemographic and economic variables across a variety of topics offers the opportunity for novel research and approaches to be undertaken

Academics from UK higher education institutions account for 90 percent of all UKDS downloads of *Understanding Society* datasets. Several types of impact can arise from using the datasets for academic inquiry, including shifting knowledge and understanding, advancing scientific methodologies, informing the development of new theories, setting and evaluating hypotheses and promoting intra and inter-disciplinary research. Academic impact from *Understanding Society* is, in part, due to the unique features of *Understanding Society*, as discussed in Chapter 4.

Defining and identifying impact is often difficult due to differences in meaning. In the REF for example, impact is interpreted as the influence the research has had. Outside of academia (for example where research has brought about a change in policy or practice); it does not include impact within academia, such as advances in methods or shifting theoretical paradigms. Impact also means different things in different disciplines and within academic work and is therefore hard to define. Further, while journal quality and citation numbers can be used as proxies for academic impact, many academics are unsure how to quantify and measure their wider contribution to societal change.

This chapter draws on evidence from the documentation review and the experiences of academics who we spoke to, to outline the range of impacts achieved. Mechanisms to achieving academic impact are also explored.

Understanding Society supports the generation of theory and advances scientific knowledge

Findings from the documentation review show that the highest volume of publications that draw on *Understanding Society* relate to health and family research. However, this evaluation has demonstrated that *Understanding Society* data is used across a broad spectrum of topics (see Table 1 in Chapter 3). Findings generated from analysis of *Understanding Society* are used in academia to create debate, establish knowledge, promote new scientific ideas and generate discussion. Nearly all academics interviewed suggest that the breadth of variables within *Understanding Society* allow for complex research questions to be answered, paving the way for strong academic impact.

The benefit of using *Understanding Society* for academic scholarship is that it "provides academic debates with more rigorous evidence" (academic user – work, education, skills & technology). Research based on such a large dataset with a plethora of variables strengthens the methodological foundation of findings. This means that researchers are more confident in their findings and the conclusions that they draw.

> *'When [a fellow researcher] tries to blow up your work and say your estimation is invalid, with Understanding Society you can say, "well I have pretty much all the information on a person that you can possibly have"'.*
>
> **Non-academic – International research institute**

The presence of variables that are strongly underpinned by theory and established definitions means that *Understanding Society* can be used to substantiate and/or validate theories.

Using *Understanding Society* promotes rigorous research methodologies to support causal explanations of findings. Our findings suggest that *Understanding Society* enables researchers to go further than descriptive cross-sectional analyses. This can help to shape and evolve the evidence base across a range of different disciplines, as highlighted in the interviews as including the link between mental health and employment, smoking and wellbeing, and financial crisis and job satisfaction. Several unique characteristics of the datasets support causal explanations. This includes the longitudinal elements and its ability to study change, large sample sizes and the diverse range of socio-demographics variables available.

> *Understanding Society is crucial to contributing to the debate about questions where there are relationships between two things that we study, causality…whether something definitely creates change in individuals.*
>
> **Academic user – Work, education, skills & technology**

Interactions between social and biological data are among some of the most recent insights that have been achieved from using the datasets. *Understanding Society* is unique in that it collects both social and biological data. The nurse biomarker and genetic data collected in Waves 2/3 includes biomedical measures of blood pressure, weight, height, waist measurement, body fat, grip strength and lung function. Evidence from the documentation review and three academic researchers shows an increasing interest in the biomarker and genetics data over time. Combining social and biological data provides academic researchers with novel ways of exploring research questions, conducting interdisciplinary research and subsequently provides new and different insights.

CASE STUDY 6: Using *Understanding Society*'s biological data to explore stress and perceptions of crime

The biological data features within *Understanding Society* have been used to explore the relationship between allostatic load (stress) and social perceptions of crime and safety of a neighbourhood. This research is helping to advance the scientific understanding of the topic by building on a body of previous research that primarily drew on self-report measures. The combination of objective biomarker data and questionnaire responses with linked crime and census data provides a novel combination of data to explore these issues. Triangulating different types of evidence helps establish more robust research, with potential impacts across the housing and health sectors. The research projects found that inequalities in the objective measurement of stress are observed in individuals regardless of their socio-economic position and neighbourhood characteristics. A higher socioeconomic position may protect against the negative impacts of poor neighbourhood perceptions.

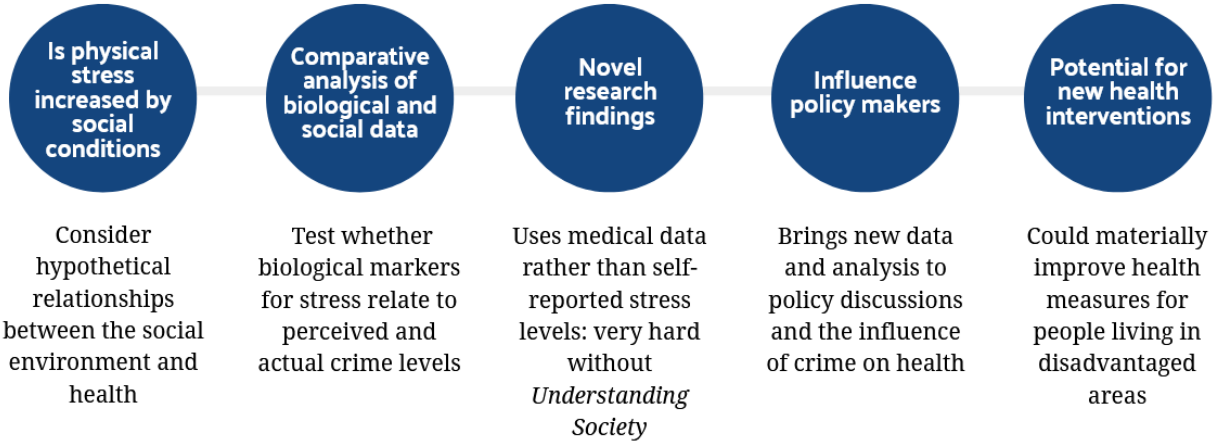


Figure 5: Analysing biological data with socio-economic data

CASE STUDY 7: Exploring testosterone characteristics with social and biological data

A team of researchers recently used biomarker data to explore the relationship between testosterone and health and social outcomes including cardiovascular health, diabetes rates and marital status. Academic impact has been achieved by supporting seven of the eight dominant hypotheses regarding testosterone levels, and these findings have been used to inform recommendations to change medical practice in relation to the time of day in which testosterone measurements are taken to account for daily variations in levels dependent on age. Testosterone measurements have only been taken at one point in *Understanding Society*. Without repeated measurements, any consideration of causality is not possible. Repeated measurements would allow for a greater understanding of the relationship between testosterone and other factors (such as weight and/or marital status) over time.

Instigating more novel research

The unique features of ***Understanding Society*** **can encourage more novel research,** meaning research which explores new connections between data, or considers existing concepts using new methods. Academically, novel research designs and findings can develop understanding in a discipline, contribute to new theory and address existing issues from a new perspective. Novel research therefore can advance understanding within a research discipline in unexpected ways, such as the research discussed below.

Using *Understanding Society* to challenge the status quo within a discipline

Research using *Understanding Society* has supported researchers to **shift paradigms relating to how social problems are framed and addressed**. For instance, public health research carried out between 2010 and 2011 that used *Understanding Society* illustrated the need to consider co-morbidities through a systemic approach, rather than individually. These findings have transformed how public health practitioners address complex health problems and inequalities.

CASE STUDY 8: Interdisciplinarity: challenging the definition of Public Health research

The unique features of *Understanding Society* dataset encourage creative thinking to address major public health issues. Interdisciplinary public health research projects, funded by the Department for Health, have explored social disadvantage and ethnicity, the interrelation of multiple risk behaviours, and the social patterning of active travel in rural and urban areas. The researchers looked at households with children within the *Understanding Society* dataset and used latent class analysis methods to illustrate patterns of individual behaviours that may be more likely to be linked to ill-health. The largest behavioural cluster was non-smoking, binge-drinking parents, who were physically active. This cluster are also more likely to be socially advantaged (white, married, with higher education and higher incomes). Using *Understanding Society* for this research offered "a much richer picture of how health behaviours are embedded in people's lives." The findings have led to a shift in how policy makers think about the communities and groups of people who engage in multiple risk behaviours, encouraging a more holistic approach of people and their lifestyles.

Academic impact can be defined as a shift in understanding and developing new theory and methods. However, these academic impacts can be achieved by non-academics who are also working to develop knowledge and shift understanding using *Understanding Society*. For instance, one organisation is currently exploring different ways of calculating household income and wealth, which they identify as having the potential to reframe discussions about benefits, employment and quality of life.

Understanding Society can help academics publish in high rated journals

***Understanding Society* can assist academics in publishing articles in highly rated journals**. The unique features of *Understanding Society* (see Chapter 4), such as the longitudinal structure with a large sample and richness of variables, enable the outputs to be well-respected in the academic community. These features increase the robustness of research and the confidence in the claims and conclusions that are made amongst peer reviewers and publishers. They also mean that advanced quantitative techniques can be employed, especially those that can establish causality from observational data; this can be a route to publishing in highly rated journals. This is a particular concern in some disciplines such as economics, psychology and health research.

> *[Journal articles that draw on Understanding Society] generally make their way to the forefront of science because Understanding Society is able to answer the sort of research questions that can't be answered elsewhere.*
>
> **Academic user – Work, education, skills & technology**

The more exposure *Understanding Society* receives through citations in journal articles, the more confidence and interest is generated in further use of the data.

Publishing quality outputs is critical to academic career progression, and for some, this is specifically associated with using *Understanding Society*. The quality of the data in *Understanding Society* means that academics feel more confident that their research will get through peer-review and be published in top journals:

> *We've got two good papers published [using Understanding Society], which has helped our careers. At the personal level it allows us to do our jobs better…because the paper was published in a top journal it's easier to convince other people that it is good, and they've been well-received.*
>
> **Academic user – Society, governance & security**

06. Social & Economic Impact

This chapter explores the social and economic impact that research using Understanding Society has achieved, showing the trajectory of impact, and exploring any potential barriers.

Summary:

- *Understanding Society* is a single, reliable source of data that can lead to a range of social and economic impacts. It is used to inform policy, inform decision-making and change business/organisational practice. However, tracking this impact can be difficult
- The robustness and representativeness of *Understanding Society* is advantageous, increasing the confidence policy makers place in findings and hence subsequent actions
- When policy makers require quick research answers, the complexity of *Understanding Society* can mean that some organisations do not have the required capacity and time to invest in the data, and other datasets may be used more readily

The social and economic impact of *Understanding Society* concerns any influence on: how society is run; changed business/organisational practices; or increased public awareness of an issue. The stakeholder testimony below demonstrates influence in policy and decision making at the highest level in the UK. At times, however, it can be challenging to attribute change specifically to the use of *Understanding Society*.

Understanding Society informs policy and practice

***Understanding Society* can be used to inform policy.** The breadth of variables and underlying social demographics within *Understanding Society* allow for a holistic representation of people's lives. For instance, some charities have used the data to show how specific metrics (e.g., low income and poverty) impacts on other areas of their lives, including mental health and housing. Others also use *Understanding Society* alongside alternative datasets or qualitative data to improve the empirical evidence for policy discussions, such as in parliamentary evidence and in roundtables with ministers. Several non-academic users felt *Understanding Society*'s robust data collection methods and high data quality (see Chapter 4) creates a stronger policy position when combined with qualitative evidence.

> *Numbers are all very well and good but if you can put that into the context of an individual's very powerful narrative about what it feels like... So, what does poverty-, how does it manifest itself? If you can put both of those together then you're much more likely to be able to change public attitudes.*
>
> **Non-academic user – Charity sector**

Earlier evidence found *Understanding Society*'s longitudinal data collection was suitable for establishing causality through quasi-experimental analysis methods. Such analyses can be used to develop and evaluate policies by comparing groups affected by the change with groups unaffected to illustrate the impact as a result of an event or policy.

> *With Understanding Society, you have a richer set of questions, and you can use these to disentangle some of the mechanisms that might explain general trends that you observe. And, if you understand the mechanisms, you can act in accordance because you can explain more.*
>
> **Academic user – Methodological Research**

There is also evidence that research using *Understanding Society* has had an impact on changing economic practice. The resulting change in economic practice is the final step in the impact journey, with the policy awareness and public understanding leading to businesses and organisations changing the way they operate in response to the evidence. This is evidenced in the vignette below relating to changes to the working week.

CASE STUDY 9: *Understanding Society* contributes to the debate to reduce the working week

A team of academics have used findings from *Understanding Society* to inform policy changes by providing evidence of a link between working hours and the impact on mental health and wellbeing. Using all waves of *Understanding Society* and the BHPS (over 70,0000 people), researchers found evidence to support reducing the standard working week, distributing surplus hours to those who need more work, and that reduced hours are 'better' for wellbeing than no work. The resulting academic paper generated worldwide interest from over 350 media outlets, including all major UK newspapers. The research fuelled a national discussion about a shorter working week in Britain, and as a result many employers are now trialling changes to assess the impact on wellbeing and business productivity. These impacts were possible through a variety of dissemination methods, including publishing videos, webinars and through the Chartered Institute of Personnel and Development (CIPD) who could feed the key findings directly into the Treasury. These findings directly fed into the Government decision to extend the furlough scheme in September 2020 to ensure people's jobs continued to be protected whilst making the scheme more flexible.

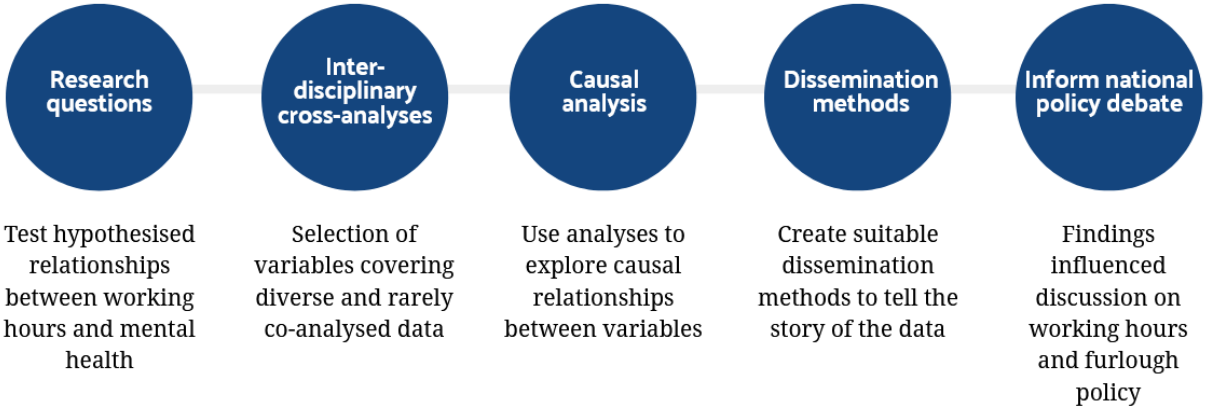


Figure 6: Combining disparate variables to influence policy

More complex modelling and analysis is possible as a result of the large number of variables and longitudinal nature of the data, as emphasised by the following vignette that explores the use of *Understanding Society* by the Bank of England.

CASE STUDY 10: Using *Understanding Society* to inform decision-making about economic stability

The Bank of England use *Understanding Society* to inform decision making about economic stability relating to consumption, spending and cyclical fluctuations in individual/household finances. The longitudinal nature of the dataset is particularly useful for the Bank to plot the distribution of savings across households with different income levels over time. The unique features of *Understanding Society* means that the Bank can develop distributional parameters in models which can help answer key research questions that are integral to the nation's economic agenda. The COVID-19 datasets are also used by the Bank to explore important employment questions, such as those relating to home working and productivity during the implementation of the furlough scheme: '*Understanding Society* is really instrumental in helping our thinking around how remote working is likely to affect productivity'. Findings from analyses of *Understanding Society* are included in published blog posts, as well as quarterly bulletin articles that feed into the Bank's decision-making committees.

Whilst identifying how research changes policy is complex, several non-academic users emphasise how data from *Understanding Society* contributes to the wider evidence base for policy discussions and decisions by providing robust analysis to inform thinking. The range of influencing factors discussed by stakeholders included ministerial priorities, political ideologies, external timing, and the wider public mood. These influences result in complex processes for making policy decisions. **Most interviewees with direct experience of policy making say *Understanding Society* is a single, reliable component amongst many sources that inform the policy decision process.** Interviewees usually feel it impossible to directly assign impact to their use of *Understanding Society,* however, it was noted across multiple interviews that this is not a problem specific to *Understanding Society* and is the case for all social surveys and research work more generally. However, many non-academic users say the quality of

Understanding Society's data is an attribute in its own right and is a persuasive tool to strengthen policy positions.

> *In terms of what goes up to a minister, although we would mention that we looked at data from Understanding Society…I don't think they even know the difference. What they want to know when you do stuff, is that it's quality assured and we've gone through a thorough analytical assurance. We're telling our ministers what we've done is robust and we believe in this analysis.*
>
> **Non-academic user, Government department**

CASE STUDY 11: Using *Understanding Society* to inform and develop the policy evidence base

Findings from *Understanding Society* help to build a robust evidence base for Ministers and policy making bodies that is used to develop future policies. The breadth and range of variables in the dataset, large sample size, and longitudinal structure enables Government departments (such as the DWP and Cabinet Office) to consider more complex research questions such as issues around COVID-19 health inequalities, mental health and wellbeing and the impact of workless households. For example, analyses of *Understanding Society* have been used to analyse vaccine sentiment and COVID-19 health inequalities experienced during the pandemic. Tracing precise impact can be challenging, but *Understanding Society* is used to inform the evidence on which decisions are made. *Understanding Society* can enable users to tell a convincing story around people's experiences which can be powerful for policy makers.

Use of *Understanding Society* to respond to policy calls

Policy questions often require quick answers which can mean that *Understanding Society* is overlooked. Most non-academic users say *Understanding Society* is a rich, complex dataset which means that it is harder to extract relevant information quickly to respond to policy requests. The barriers they cite include the longitudinal structure, although the most recent cross-sectional wave is sometimes used for exploratory or policy queries. The annual data collection cycle can mean data is not as relevant as some users would like. Others plan their analysis to get the most out of the dataset. Planning requires time which is often unavailable in a policy context, especially where the analytical need is reacting to policy or lobbying towards a current issue.

> *I think one of the areas where we might be able to justify the time would be the state pension age review? I think that's somewhere where we've got enough warning, long enough in advance, that we're going to be influencing on something. And it's an important enough topic to us that we can do some proper, rich research and analysis. Often, we're very reactive. So, government will make an announcement and we need to campaign against it or something, and it's now. So you can't do a meaty, proper, detailed analysis [of Understanding Society], we just need to start shouting.*
>
> **Non-academic user – Charity sector**

The fast-paced environment of public debate means that policy needs can change quickly and move on significantly over the time a research project is conducted. The policy requirements by the time of publication can move on rendering research less useful.

> *When you start [policy-facing research], it can be highly policy relevant, by the time you report, it has no traction at all.*
>
> **Academic user – Health & human behaviour**

***Understanding Society* is well-placed to inform policy development.** Several non-academics feel *Understanding Society*'s policy strengths were informing policy design and government decision-making, particularly when government needs to fine tune a policy. These interviewees feel exploratory analysis using *Understanding Society* to test hypothetical questions offer greater potential impact compared to the reactive analyses discussed above.

> *Things that tend to have more impact tend to come from policy directly…if it comes from them and you can see it's a clear ask of "we need this because we need to make a decision in relation to x policy"…when the questions are coming from policy directly, it's great because it means they're planning to use it and they will bring it to the minister's attention.*
>
> **Non-academic user – Government department**

The *Understanding Society* COVID-19 waves were frequently used for policy impact. The COVID-19 dataset ran between April 2020–September 2021 and covered the changing impact of the pandemic on the welfare of individuals, families, and communities in the UK. Most non-academic interviews said increasing the frequency of data collection and integrating COVID questions into *Understanding Society*'s breadth of lifestyle measures played an important role in setting COVID policy. The richness of the dataset, coupled with the regular data collection meant that a government department could combine social justice policy, food poverty policy and education policy to inform a policy decision regarding food parcels during COVID-19:

> *There was a lot of policy making going on around food during the pandemic. They did these food boxes and things like that. So, yes, [Understanding Society] definitely influenced policy in that way.*
>
> **Non-academic user – UK or devolved government**

Direct policy and economic impacts are not always immediately clear

Evidence using *Understanding Society* contributes to "feeding the [policy] machine" but **pinpointing the extent of the impact is difficult**. For many non-academic organisations, producing evidence that informs political debate is an important metric of

'success', irrespective of whether their findings are specifically referenced in a green/white paper. For third sector organisations, they see their work as contributing to the debate to change policy and bring about better conditions for individuals; if change occurs, this is classed as success. In both cases, *Understanding Society* in combination with other evidence contribute to impact, but the precise contribution of *Understanding Society* is unknown.

> *Sometimes the conversation just shifts and things happen that you have been calling for for a decade. [The actual impact] might not be clear to you, or clear to you only. When there are small policy wins, we just go 'well it's great that that happened.' But we often don't know if we had anything to do with it.*
>
> **Non-academic user – Think tank**

Policy and economic impact require strong networks of influence

Organisations that use *Understanding Society* to achieve societal and economic impact have strong links to government and policy makers which enables their research to quickly enter the policy making system and be considered. For charities and other third sector organisations, having these networks can be more difficult, but combining with other organisations or groups to lobby or influence policy can help increase their influence. Large charities can have a strong network of influence and can use their national platform to highlight the key issues that are important to their cause. As the following example shows, *Understanding Society* can play a key role in providing key statistics to support the charity's work.

CASE STUDY 12: Third Sector organisation uses *Understanding Society* to campaign for policy change

Age UK uses *Understanding Society* for fundraising campaigns and policy-informing research. Analyses using the dataset are used by the charity, alongside findings from other datasets, to highlight key issues to campaigners, MPs, civil servants, and the public. These statistics help to keep key issues relating to older people a priority and relevant to national and policy-maker conversations. The different waves and longitudinal design of *Understanding Society* allows the charity to explore how life events can influence a person. For example, Age UK has used the study to examine the effects of bereavement on loneliness. The richness of variables is one of the most powerful features of the dataset, which has enabled Age UK to create an Index of Wellbeing in Later Life. This provides policy makers and the general public with information about what is important for older people's wellbeing. The inclusion of younger people in the *Understanding Society* sample is also valuable and allows the charity to explore caring responsibilities across the population.

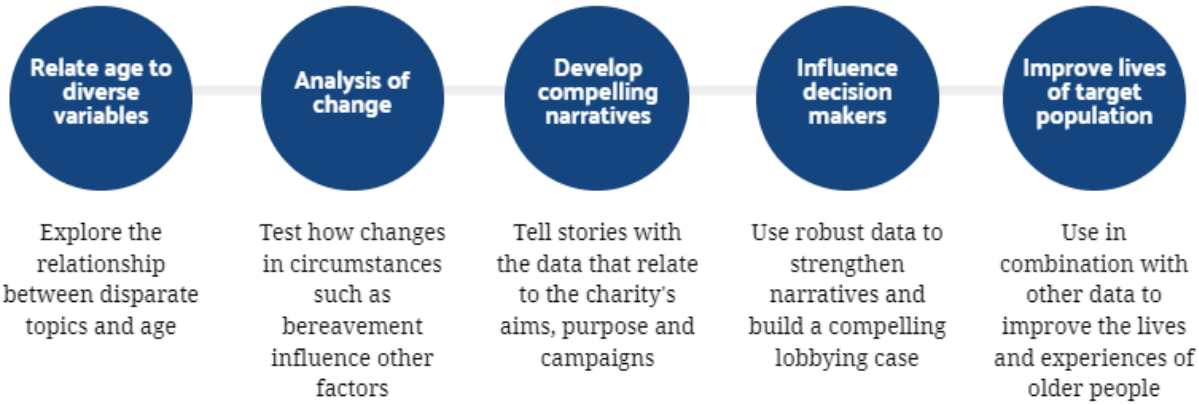


Figure 7: Using *Understanding Society* for charitable aims

Factors that affect the ability to influence policy and practice

Low sample sizes for devolved nations in *Understanding Society* limits its use for policy making in Scotland and Wales[7]. When conducting analysis of the devolved nations, there are often smaller sample sizes, particularly when conducting sub-group analysis. Whilst interviewees referred to this as underrepresentation, the actual sample issued per head of population is larger for Wales than both Scotland and England[8]. As a result of the smaller absolute sample sizes in devolved nations, there are limitations with what *Understanding Society* can be used for:

> *As a devolved nation, it's not always necessarily as clear with stuff that goes across the UK, whether or not [Understanding Society] will be able to answer [devolved nation] specific questions… sometimes, if there's a call for evidence, and I bring them stuff that's English or even UK level, they'll say, 'No, we don't want that. We need something [specific to this country]'.*
>
> **Non-academic user – Devolved government**

Ideas and theories that are generated through *Understanding Society* can help challenge dominant ways of thinking about policy. Shifts in understanding can affect policy and how policy makers approach decision making, as exemplified by this academic user:

> *The idea of knowledge translation, what goes from research to policy, is pretty narrowly conceptualised as data and results. But actually, what travels are our ideas or*

[7] No interviews were undertaken with members of the Northern Ireland devolved administration who power-sharing arrangements are in abeyance

[8] For Wave 11, the total sample issued to eligible adults relative to the total population of each nation (all per 2,000) is: England 1.22; Wales 1.74, Scotland 1.29 and Northern Ireland 2.85.

> *ways of thinking and frameworks… I think evidence in and of itself doesn't change policy. The ideas and ways of understanding that it generates do.*
>
> **Academic user – Health & human behaviour**

Their experience shows that research using *Understanding Society* has the impact of changing the paradigm of how policy people think about public health issues, shifting the way policy is thought about and decisions are made.

07. Factors that limit using *Understanding Society*

This chapter explores perceived limitations of Understanding Society and the consequences for achieving impact through applied research. Across all stakeholder engagement activities, the positives and impacts of the dataset were extensive, with any critique couched as an aside and with far less weight attributed.

Summary:

- The complex structure and design of *Understanding Society* can be a barrier for new and inexperienced users. A considerable investment of time is needed to become a proficient user of the datasets, however, once competency is reached, users are effusive in their praise and use of *Understanding Society*
- Outdated, dropped, skipped or modified variables are problematic for users that carry out longitudinal analysis
- The resource and timescales required to access the Special License Data can deter people from using these datasets
- Whilst an advantage of *Understanding Society* is that it enables sub-group analysis, depending on which sub-groups and intersectional selections are chosen, small sub-group sample sizes and high attrition rates are a concern for some users. This is particularly the case when combining demographic characteristics such as ethnicity with age/gender/location and the linked datasets with low matching rates

Despite the numerous advantages and unique features that *Understanding Society* offers, interviewees highlight several limitations. These can impede their work and the potential impact that can be achieved. As highlighted in Chapter 4, some of the unique features of *Understanding Society* are viewed as both strengths and limitations, often depending on the level of experience of the user – with inexperienced or new users citing barriers and limitations as more problematic compared to experienced users. It is also the case that non-academic users experience greater initial barriers to accessing the data and appear to cite more limitations than academic users. The limitations reviewed in this chapter were identified across interviews, but the relative strength of these limitations is lower than the advantages that *Understanding Society* offers, as depicted in Figure 3.

Limitations discussed in this chapter include issues around the complexity of the data, the content and structure of *Understanding Society* and accessibility of the datasets.

Interviewees have also provided some areas of consideration to further enhance the user experience of *Understanding Society*.

Understanding Society's complexity

***Understanding Society* is complex and can be difficult to use for inexperienced and new users.** Just over half of the interviewees say that it is challenging to learn how to use *Understanding Society* and become familiar with the datasets. A significant initial time investment is needed to realise the potential benefits of the datasets, but this is not surprising given the richness and breadth of the data.

> *You can run a survey, right? That's a much cheaper and easier way to tick the research box than learning how to use Understanding Society.*
>
> **Academic user – Society, governance & security**

The context for this quote was exploring the differential use between academics and non-academics, suggesting that a bespoke survey to address a key policy question may be more effective for small think tanks or policy units than upskilling to use *Understanding Society*. Users therefore expressed an assumption that there is a high level of quantitative skill required to use the dataset, a skillset that may not be as commonly found outside of academia – "your average policy analyst in a think tank doesn't have a PhD in economics" (academic user – Society, governance & security)

Consequently, non-academic organisations highlight that the upfront time investment needed to learn how to navigate the datasets limits their engagement. These organisations often commission research externally or overlook the dataset in favour of less complex datasets. For one charity, they commission analysis when the research is too complex to be done in-house, and often these pieces of research involve *Understanding Society* "because it's a very large dataset and you need good experience using statistical software to use it in the way that we want it to be used" (non-academic user – Charity sector).

However, the **time-investment required to learn how to use it is accepted** as a necessary part of then having access to such rich data.

> *It's such a wonderful, long-term, longitudinal data set – but you can't just start using it in an hour.*
>
> **Academic user - Society, governance & security**

Occasional users find it harder to use *Understanding Society* data - Users need to gain a 'threshold' of experience to become confident and proficient in using the datasets. Some users suggested it can take three months to learn how to use the dataset sufficiently to produce meaningful analysis. Engaging with the available resources and training helps

users to meet this threshold. In turn, the value of the datasets to their work increases significantly:

> *It's not something that you do casually... You need to decide you're going to invest a substantial amount of time in getting to grips with it before you can actually do anything with it.*
>
> **Academic user - Society, governance and security**

Working in collaborative teams and drawing on varied experiences to extract meaning and analysis from the data can help to overcome issues of complexity. An example of a collaborative project explored testosterone characteristics, and within the research team, the interviewee provided subject-specific content knowledge and worked with other researchers who were more familiar with *Understanding Society*:

> *If I had done this project by myself it would have involved a considerable investment in time to learn the dataset adequately to be confidently running analysis on it. They knew the datasets very well, but they didn't know much about testosterone. So, it was a very nice melding of forces and that is, you could say, I provided the knowledge background of the substantive area and they, being very familiar with the dataset itself, they did almost all of the runs and initial analysis.*
>
> **Academic user – Health & human behaviour**

There are some specific barriers for new users. In particular, these relate to navigating through the different files, waves and datasets, including how to link data together. For new users, the complexity of the dataset can be intimidating and confusing to understand which files and variables they need to address their specific research questions.

Irrespective of level of experience, a commonly identified feature of the dataset that is confusing for users relates to the multiple weighting variables and when these should be applied. Some interviewees express the complexity and number of the weighting variables can make them question whether they applied the right weightings and hence whether their analysis is correct.

> *It's a bit of a shocker when you've got your variables in front of you and you type 'weight' and you can't decide between 20 different options.*
>
> **Non-academic user – Independent researcher**

A couple of interviewees feel some users may inadvertently incorrectly apply weightings in analysis and take shortcuts due to the complexity. This is partially because it is so complicated, but also due to the time required to learn it effectively.

> *One of the things I consistently struggle with the most is the weighting and when you're just doing a couple of waves in which you have the variables that you need, if you're*

> *skipping other waves, making sure you've got the right weights assigned is one of the biggest headaches and one of the biggest things to prevent people from using the dataset.*
>
> **Non-academic user - Think tank**

Content and structure of *Understanding Society*

Sample size and attrition rates

Aspects of the content and structure of the datasets can also present barriers that prevent people using the data for their research. The total cross-sectional sample size in *Understanding Society* is relatively large (in the context of longitudinal household surveys); however, for some research questions there can still be an issue with **small sub-group sample sizes**, particularly for intersectional research that requires data analysis at smaller geographical levels, or of different ethnic groups. Sample sizes for research using devolved nations is also often too small to do meaningful analysis. Across interviews with researchers from the devolved governments (n=2), plus interviews with researchers working on samples based in Scotland and Wales (n=2), there is a perception that there are not enough Scottish or Welsh people in the sample. This interviewee highlights the limitations of low sub-group sample sizes for conducting intersectional analysis in their research:

> *When you break it down by 16 to 24-year-olds who are Asian in London, that's three different things, because you're breaking down the UK sample into London, you're breaking down to Asian, breaking it down into 16 to 24, there's hardly anybody there, so you can't be confident about the result.*
>
> **Non-academic user - Local government**

Consequently, interviewees raising this point think more powerful analytical models cannot be used on smaller sub-samples which can lead to uncertainty about the reliability and validity of findings. This may influence choices to publish research and hence access to potentially useful new knowledge. Planned intersectional analysis can lose statistical power and become meaningless. However, interviewees are also keen to emphasise that *Understanding Society* provides access to one of the largest and richest samples of household and individual level data. Issues with sub-group sample size and attrition are not unique to *Understanding Society*. Despite issues of sub-sample sizes, approximately a quarter of interviewees expressed a note of caution that **any increase in sample boosts may come with a compromise to other valued parts of the dataset, such as the depth of variables available**.

Attrition rates and the characteristics of dropouts are often different to those that remain within the sample. Interviewees identified that the groups with whom drop out or non-response may be particularly problematic include students, parents with young children (during COVID data collection), and individuals with severe mental or

physical health or financial difficulties. Underrepresentation of these groups risks unrepresentative findings and may result in biased interpretations. Interviewees also expressed that attrition rates are particularly problematic for later waves of the COVID-19 dataset and the Youth Data files.

> *One of the problems with the COVID waves is that there's quite substantial attrition and spottiness across the waves… Looking at the profile of the people who had dropped out, had answered the COVID wave 1, but then not any others. Dramatically different in certain ways, people with little kids, just out of the picture.*
>
> **Non-academic user – Independent research institute**

Despite issues with attrition, users of the COVID-19 dataset are typically effusive about its timeliness, longitudinal analytical possibilities, regularity and richness. Interviewees were also cognisant of the inherent tensions between ideally wanting more regular data collection points and the associated impact that this can have on attrition and response rates.

The mode of survey delivery can cause issues with sampling and representativeness. Users acknowledged the data collection mode necessarily changed because of government COVID-19 regulations preventing face-to-face research. A few interviewees (both academic and non-academic) note that "*the shift to online [has meant] there's a large digitally excluded community who we are not getting*"[9] (academic user – health & human behaviour). A potentially less representative sample could weaken conclusions about society.

Sample sizes are also particularly problematic for users of the NPD linked datasets. Interviewees perceive the ability to use linked data as a positive aspect of *Understanding Society*, but **small samples sizes of linked cases can prevent meaningful analysis**.

> *I abandoned using the NPD data because there were not enough cases. The documentation said that there was this many cases, but when you looked, linkage was only possible for a small proportion of that.*
>
> **Academic user (focus group) – Health & human behaviour**

[9] This is the perception of the interviewee – as part of both the mainstage annual survey and the COVID-19 survey, data was collected using both online and telephone methods.

Consent to match to education records was obtained in Waves 1 and 4, and on a regular basis for the data that is already deposited at UKDS[10]. This means that it is only possible to match records for those in these two waves, resulting in a smaller than ideal sample size. Matching has therefore happened on two occasions (data up to 2012/13 and 2017/2018). Whilst there are intentions to link data again, users of the NPD data would like more frequent linking (annually) to reflect the ongoing educational changes that individuals go through during their lives.

> *The linking needs to be an ongoing where every year they refresh it, because everyone is ageing, people are acquiring qualifications. Refreshed educational records every year would allow us to actually use it to look at educational questions in a much more thorough way.*
>
> **Academic user – Work, education, skills & technology**

Variables and data types

Nearly half of all interviewees identify that some variables in the study have limitations. Some think variables are too general (e.g., information on financial variables in *Understanding Society* is more general than in specific financial surveys) whereas others feel the number of derived variables is insufficient meaning that their variables of interest require further manipulation by the user before they can be used in analysis. Several interviewees say there are inconsistent definitions for variables.

Outdated variables may not reflect the current context. Some variables within *Understanding Society* may not be appropriate or timely enough for users. The understanding and use of some social research topics has changed dramatically in recent years. For instance, awareness of gender identities and sexuality have increased and society is increasingly less structured on a binary gender classification[11]. Some feel *Understanding Society*'s design should reconsider how gender is approached to ensure the spectrum of identities are included. A further example of where developments in the discipline may be moving quicker than the rotation of questions in *Understanding Society* is research into climate change and modes of transport[12].

> *We had questions on fuels and types of fuels that you could run your car on and, in three years, there might be a couple of new things and maybe we won't have diesel*

[10] For information on the consent and matching of education records see: https://www.understandingsociety.ac.uk/sites/default/files/downloads/documentation/user-guides/mainstage/7642_user_guide.pdf; https://www.understandingsociety.ac.uk/sites/default/files/downloads/general/long-term-content-plan.pdf

[11] *Understanding Society* has implemented this suggestion in their questionnaires, however, at the time of interviews and writing, it has not fed through into the data available publicly.

[12] *Understanding Society* are also conducting a review of the environmental content, thus potentially addressing these issues identified by this interview quotation: https://www.understandingsociety.ac.uk/2022/04/08/help-shape-our-questions-on-environmental-attitudes

> *anymore, and if the question has been set three years before it just won't be accurate. There's a long gap between thinking through a policy area and finalising a survey.*
>
> **Non-academic user – Government department**

Further areas in which questions can become quickly out of date include IT and digital technology. With these examples, having variables that are not as accurate to the current context means that analysis cannot accurately reflect what is happening in society. Any conclusions will therefore have less policy or academic relevance.

A lack of derived variables can make the dataset less accessible for some users. Absent derived variables means that users must undertake more steps in their analysis which can be a barrier or introduce uncertainty. Users have suggested that derived variables would be useful for data about employment history and income[13]. Other users are confused by the way in which derived variables are created and would require more information on how the variable is generated[14].

> *It would be nice if you could just click on a link that said, 'This derived variable is a combination of these other variables and this is the formula or algorithm that they've used to weigh all the other variables to get this one.*
>
> **Academic user – Society, governance & security**

Some variables are difficult to interpret. Across the broad spectrum of variables included in *Understanding Society*, a small proportion of users specified that choosing the correct variable was difficult. For instance, variables such as age and sex that have corrected versions can cause confusion as to which variable is correct, leading to uncertainty about what individual questions mean.

Dropped, skipped or changed variables can be problematic for longitudinal analysis. Many variables within *Understanding Society* remain consistent across multiple waves and are not affected by new/changed definitions, as mentioned above. Having regular and uninterrupted measurement of these variables is important for longitudinal analysis. The rotating question structure can be frustrating for some users when their variables of interest suddenly are dropped or changed. Examples of this from interviewee experiences include changing how health conditions are recorded, the number of hours of childcare question not being included in recent waves, or specific combinations of variables only being present in a couple of waves (e.g., satisfaction with NHS and NHS

[13] *Understanding Society* are currently working to produce employment history data and other event history data which will be available for users to access

[14] *Understanding Society* provide syntax for some derived variables and are working for doing this for all variables.

usage). This disrupts their data series for longitudinal analysis and prevents a thorough analysis of change and potentially causal explanations, as highlighted by this interviewee:

> *There are sometimes some odd things where they skip a variable in one wave and you're like, why have you skipped this variable? That's a bit of a pain.*
>
> **Non-academic user – Government department**

Interrupted datasets can impact on the extent of statistical analysis that is possible, because the more consistent variables within a dataset means that more powerful statistics can be used.

Insufficient biomarker data can prevent causal inferences being made. Biomarker data was collected in Waves 2/3 (of the General Population Sample and the BHPS sample) during nurse visits. During these nurse visits, bio-medical measurements were collected, including measurements of blood pressure, weight, height, waist circumference, body fat, grip strength and lung function, as well as blood samples to measure blood fats, diabetes, inflammation, anaemia, liver/kidney function and hormone (e.g., testosterone) levels.

Currently, there has only been one nurse visit and associated data. Users of this data all identify that regular and/or repeated biomarker data collection would allow users to make stronger causal inferences. More frequent biomarker data collection points would enable longitudinal analysis and enhance linked biological and social data. This is explored more in Case Study 7; whereby more frequent testosterone measurement would enable causal predictions to be established:

> *The critical variable for us was testosterone, and unfortunately that was only measured in one wave of Understanding Society, unless testosterone is measured again, there's no further use [of Understanding Society] to us.*
>
> **Academic user – Health & human behaviour**

Accessing *Understanding Society*

Special License files

There are three categories of access to *Understanding Society* data, depending on the likelihood and potential risk of disclosure.

- Most users access *Understanding Society* data via the End User Licence.
- More specific data (e.g., main survey data with detailed classifications – month of birth, local authority district or parliamentary constituency) can be accessed via the special licence; and
- More disclosive data is provided through Secure Access (e.g., full dates of birth, NPD data, grid reference geographical data).

Eighteen interviewees applied for special licence access, with 6 interviewees having accessed data (usually full grid reference from which postcodes can be identified) via the secure access route. Gaining higher levels of access allows for more detailed contextual information to be used within analyses, including specific geographies and income data.

Perceived barriers associated with the secure and special licence files can have a **negative impact on the types of research that can be conducted.** Despite the benefits of accessing more detailed data, users with time constraints (often non-academics with policy pressures or PhD students with deadlines) are put off by the bureaucracy of applying and choose instead to use the end-user licence.

> *The special license access is a constant challenge, it takes 2 to 3 months to get approved for access through the UKDS… if I have new person coming in in September and we have to produce something by the end of March, it can be a real challenge because you can't access the data for several months until you've got approval.*
>
> **Non-academic user – Government department**

As a result of the challenges associated with obtaining more secure levels of access (and detailed data), researchers may be more conservative and cautious with the research they choose to do. The consequence of this is that there can be a limit to the development of methodological impact and capacity within the research community. As one academic user summarises:

> *What actually happens in practice is that you refine your question to avoid using sometimes some of the secure information if you don't have to but the obvious loss is that you tend to produce more of the same sorts of research in that, you might go for the safer and easier to do piece of research rather than the slightly more cutting edge piece of research because of the kind of constraints that you face in trying to do it.*
>
> **Academic user – Work, education, skills & technology**

International users cannot access the secure license files and need to use UK-based collaborators to access the data they needed[15]. This may limit international collaborations and comparisons, reducing the perceived value of *Understanding Society* internationally. On the other hand, controlling this access could also encourage collaboration through necessity, as collaboration may be the only way to access data.

[15] As previously mentioned, secure access is now possible for researchers in Germany, France and the Netherlands.

> *I think, one of the negative aspects, that you need to be based in the UK to access some types of data and if you are not in the UK you can't, or you need to collaborate with someone there*
>
> **Academic user – Methodological research**

Enhancing the user experience

Areas that interviewees highlighted that the ESRC may wish to consider to further enhance the user experience of *Understanding Society* are largely focused on streamlining processes to reduce the time and cost investment required by users. Some of the suggestions below may already have been considered or implemented, but these are identified by participants as areas to consider.

Ensuring supporting documentation is more user friendly could be achieved by including short worked up practical examples that outline the steps involved in carrying out longitudinal analysis. Interviewees suggested a weighting matrix table to illustrate when different weights should be applied would be helpful. There may also be benefits gained in sharing code that users can apply to append and merge different waves of the datasets, together with a short document that outlines the different types of data linkage that can be carried out[16]. International users voiced that they would find it helpful to have access to a summary table that outlines data availability for international analysis and what licences are required. Low awareness and use of the Innovation Panel could be improved with better documentation and resources relating to it.

Clear and timely communication outlining when variables are removed or changed would enhance the user experience by alerting them to when the changes will take place. This will enable users to carry out necessary revisions to their research schedules. Users would also appreciate greater consultation about planned changes to variables.

Increasing traffic to, and engagement with the *Understanding Society* website could be achieved by reviewing the recently revised layout and improving signposting to relevant links and information. Clearer positioning of the Policy & Partnerships Unit on the website would help increase its perceived importance and awareness to users.

[16] Within the *Understanding Society* documentation, a table and video are provided regarding how to select weights, and code is provided on merging different datasets. However, the presence of these quotes suggest that signposting to the different content may need to be clearer to ensure all users are aware of the complement of support available.

Re-introducing support via initiatives such as the previous charity voucher scheme would support non-academic and third sector organisations to test new research questions and methods in the knowledge that analytical support is available to them.

08. Conclusions

In this final chapter, we return to addressing the overarching research questions before summarising the key thoughts of stakeholders towards Understanding Society.

What do stakeholders think of *Understanding Society*?

All stakeholders and users of *Understanding Society* consulted during this evaluation were overwhelmingly positive about the dataset, what it offers, and the impact that research drawing on the data can achieve. Users consider the dataset to have a brand name with a global reputation which is viewed as a 'great national treasure'.

The overarching sentiment conveyed by users of *Understanding Society* is that the dataset is a worthwhile, valuable and important resource that is the 'envy of the world'. It has many unique features and key strengths including household-level and individual occupant samples, a longitudinal structure, numerous and rich variables available for analysis, large sample sizes and timely data. These unique features mean that users view the dataset as a robust and 'trusted' resource. Major changes to the structure of the dataset would risk adversely affecting the extent to which impact could be achieved. For most interviewees, *Understanding Society's* many strengths outweigh the few perceived weaknesses. A minority of users, in particularly infrequent and non-users think the complexity of the dataset can make it more challenging to carry out analysis.

How is *Understanding Society* used by stakeholders?

Most consulted stakeholders used the main individual/household dataset and the COVID-19 files. Most conduct their own primary analyses, either consisting of cross-sectional or longitudinal analysis. More experienced researchers typically exploit the longitudinal features to demonstrate change over time and explore causal relationships between variables. Primary research and longitudinal analysis of *Understanding Society* is believed to offer the most benefits and potential pathways to impact for users. *Understanding Society* data is frequently used in PhD research projects, developing quantitative analytical capacity, thus ensuring the pipeline of quantitative researchers both in the UK and internationally, remains strong.

A small subset of users engage with *Understanding Society's* innovation panel to test novel methodologies and apply more complex analytical techniques. Users frequently engage with the suite of training materials provided by *Understanding Society* to show them how to get the most from the data. This helps users to become familiar with the data, thereby increasing their skills and competences and improving the quality of research outputs. This in turn enables research findings derived from the study to influence local and national policy making decisions.

What have stakeholders been able to achieve which would have been difficult and/or impossible if *Understanding Society* did not exist?

Interviewees think the UK's data infrastructure would be significantly weaker without *Understanding Society*. There are no other datasets as large and rich which provide longitudinal lifestyle data on individuals and households. The longitudinal structure facilitates policy development and allow users to evaluate the impact of policy change on people's lives:

> *If we think about why policy makers introduce reform – it's about producing a change in behaviour. By having that longitudinal dimension [in Understanding Society], we can follow individuals over time and we can actually answer the question that we need to know [about policy impact].*
>
> **Academic – Health & human behaviour**

Users emphasise that without *Understanding Society* they would still be able to carry out their research, but the quality would be weaker. Many would have to change their research question to fit the (limited) data or change methods of analysis. These findings collectively suggest lower confidence in findings from other sources which may generate less impact on policy and societal decision making. This is exemplified in the Department for Work and Pension's case study where their work on the impact of workless households on children would not have been possible without *Understanding Society* because the variables of interest are not available in existing administrative datasets.

What social benefits and impacts has *Understanding Society* underpinned, and how?

Research using *Understanding Society's* expansive data delivers academic, and social and economic impact. The robust structure of the dataset and the richness of information creates a platform for rigorous and impactful research. Stakeholders cite a range of academic impacts including the development of new theories and knowledge insights, and linking biological and social data in novel ways. This in turn can advance scientific knowledge and change paradigms regarding how social problems are framed.

Social and economic impacts generated from using the dataset are far-reaching and include informing government policy, supporting changes to business working practices, feeding into parliamentary round tables, and informing government department papers for ministers. Key examples evidenced in the Case Studies (Appendix A) include informing the UK Government's review of social distancing in response to COVID-19, informing a UK-wide trial of a four-day working week, and use by national charities to inform their campaigns and educate policy makers and wider public about their cause.

Data and research findings drawing on *Understanding Society* initially form part of the wider evidence base that is used to inform policy. However, measuring the direct pathway to policy impact from using *Understanding Society* can be challenging. Despite this, users view the dataset as a valuable tool that is often considered by policy makers in the UK.

Implications of this evaluation of *Understanding Society*?

In addressing the key research questions, this evaluation shows that stakeholders consider the dataset essential for their work, as well as having wider social and economic benefits. Continued support for the dataset is important to ensure that these far-reaching impacts can be achieved. As one researcher emphasised:

> *I hope [Understanding Society] continues. I hope it doesn't get thinned out or anything because it's a very valuable resource that's good for the economy, good for the country, and good for research.*
>
> **Academic user – Society, governance & security**

Understanding Society provides users with a cost-effective way to carry out cutting-edge research. In the absence of the study, users note that with independent funding, they could carry out their own bespoke cross-sectional surveys. However, this would be on a smaller scale and the depth, breadth, and the longitudinal aspect would be near impossible to recreate; this would severely hamper the potential for impact.

The cost of *Understanding Society* could be reduced by simplifying the structure of the dataset, reducing the sample size, including fewer boost samples, and/or reducing the number of variables. However, the consequences on the quality of research and policy impact would be significant. Continuing the study in its current form offers researchers unrivalled access to information on people's lives and experiences, and how these change over time. Reducing *Understanding Society's* scope and coverage would limit opportunities for cross-disciplinary analysis. Smaller sample sizes would limit the trust current users place in findings from the data and increase error margins, especially where the frequency and size of boost samples were also reduced. Researchers, and those who use research based on *Understanding Society's* data, would be less confident about the conclusions reached.

Without *Understanding Society*, the quality of UK social science research would diminish; in some cases research ideas would remain unexplored and research questions left unanswered. Data from *Understanding Society* can answer comprehensive and complex questions, and in its absence the evidence base for exploring questions of interest is partial. This means that any research outputs will be weaker. Therefore, *Understanding Society* strengthens the evidence base for scientific enquiry into UK social issues:

> *We wouldn't be able to fully answer our research questions. It would be like writing half a paper. It wouldn't make any sense, to some extent. If you went to a parliamentary committee, you could only give partial answers, which isn't really sufficient. We wouldn't be able to give comprehensive answers because we wouldn't know them. It would be like starting a painting and only doing half of it.*
>
> **Academic user (focus group) – Health & human behaviour**

Appendix A: Case Studies

Case Study	Content
1 - Providing evidence on family and household behaviours to inform Government COVID-19 policies	Research by the ONS used *Understanding Society* to help inform UK Government policy during COVID-19.
2 - Using *Understanding Society* to build quantitative skills and increase productivity	Due to the high value placed on the dataset, the IFS have created an extractor software which allows their analysts to become quickly familiar with the *Understanding Society* data and more confident in their analysis.
3 - The Innovation Panel to foster methodological innovation	Example of the use of the innovation panel for changing the paradigm about successful ageing and methods for exploring this empirically.
4 - *Understanding Society* meets the shared needs of research and policy users for research into transport	Explores the ways in which academic researchers and Department for Transport analysts have used *Understanding Society*, showing how the different variables can meet the needs of different user types.
5 - The validation and testing of a forecasting framework using *Understanding Society* to develop innovative methods and advance statistical methodologies	*Understanding Society* has been used to test a model that forecasts future life events for US military personnel, demonstrating innovative methodologies and advancing scientific understanding.
6 - Using *Understanding Society's* biological data to explore stress	Researchers have used *Understanding Society* to explore the link between allostatic load (stress) with environmental and crime data.
7 - Using biomarker data to gain new interdisciplinary insights into the interaction between testosterone and social outcomes	Researchers have used *Understanding Society* to test hypotheses about the levels of testosterone in men.
8 - Interdisciplinarity: challenging the definition of Public Health research	Using *Understanding Society* research to shift understandings of Public Health and change how policy makers consider health behaviours.
9 - *Understanding Society* contributes to the debate to reduce the working week	Research by academics has shown that shorter working weeks can improve employee mental health which has influenced policy and practice across the UK.
10 - Using *Understanding Society* to inform decision-making about economic stability	The role of *Understanding Society* in assisting the Bank of England in understanding evidence to inform decision making to ensure economic stability.
11 - Using *Understanding Society* to inform and develop the policy evidence base	Evidence of the Cabinet Office and Department of Work & Pensions that use *Understanding Society* to inform the evidence base for policy.
12 - Third Sector organisation uses *Understanding Society* to campaign for policy change	Age UK's use of *Understanding Society* to campaign and contribute for calls to change policies relating to older adults.

Case Study 1: Providing evidence on family and household behaviours to inform Government COVID-19 policies

The *Understanding Society* COVID-19 survey has enabled researchers to provide timely, evidence informed findings to help the Government answer a range of policy related questions. In response to the unfolding COVID-19 pandemic, *Understanding Society* were quick to respond and adapted their data collection process to include a COVID-19 survey, for which data was collected in monthly waves. The COVID-19 survey contains specific questions about the social impacts of the pandemic on individual and household behaviours. This timely and novel data is not available in any other survey. COVID-19 survey questions can be linked to the main survey, which has the benefit of providing a pre-COVID-19 baseline from which to conduct analyses.

During the pandemic, researchers from the Office for National Statistics (ONS) specifically used *Understanding Society* including the COVID-19 datasets to provide insights into household relationships. These relationships included how far families travelled to see each other; travel times to relatives; whether grandparents lived with families; the distances between nearest relatives for people who live alone; and, living arrangements of cohabiting couples. This research had wider societal impacts by providing an evidence base to the Government to inform their decisions on further COVID-19 restrictions during 2020:

> *The distance from nearest relative was timed to enable data to be included in the evidence used to inform decisions around Christmas 2020 and the lockdown restrictions.*

The ONS has used *Understanding Society* to inform a number of outputs that have been used in policy decision making. For instance, drawing on the COVID-19 survey informed a review of the social impact of COVID-19 on ethnic groups, and further data from the COVID-19 survey was one of several sources submitted to the Government as part of the social distancing review. Analysis using *Understanding Society* also informed part of an evidence pack about the impact of social bubbles on different age groups and on keeping schools open. ONS analysts stress how *Understanding Society* data was not the only source of data that informed the final policy decisions but was a significant contributor when presenting data to government.

Understanding Society offers several unique features for analysts using the COVID-19 dataset. Of particular value is the ability to track changes in individual and household behaviours over time, through linking individuals in the data to their non-COVID-19 wave data. The depth of data on relationships outside the family unit is also a key strength. These unique features supported the ONS to demonstrate societal impact through identifying changes to family behaviours since the start of the pandemic and predicting what behaviours might change if the Government implemented social restrictions.

Case Study 2: Using *Understanding Society* to build quantitative skills and increase productivity

The complex nature of the *Understanding Society* dataset means it takes time to develop the necessary knowledge and skills to proficiently use it. To build capacity and increase productivity when using *Understanding Society*, the Institute for Fiscal Studies (IFS) have created their own extraction software. The software is particularly useful in supporting researchers who use the dataset, to understand household income and spending behaviours. The tool has been developed because data from *Understanding Society* is valuable to their work, and the organisation wants to ensure they can access and use the data effectively.

The IFS have used *Understanding Society* for numerous projects, primarily using the Individual and Household main datasets. The longitudinal and household structure of the data is particularly beneficial for their research. It enables them to link parental and child data to explore longitudinally, how parents' wealth correlates with their children's wealth and other outcomes (e.g., home ownership).

The IFS have created their own *Understanding Society* "extractor" to ensure staff at can access *Understanding Society* to familiarise themselves with it and carry out analysis efficiently:

> *People who are new to the survey can just run some code which prepares a data set for them that is easier to start working on than the raw data.*

Creating this extracted dataset means the code for the derived variables can be shared, which helps to bolster quantitative capacity among researchers quickly. Using an extracted dataset with consistently defined variables commonly used by the IFS also has time saving benefits, thereby increasing the quality of outputs and productivity:

> *We want researchers to be able to extract ready-to-use data sets that are consistent and have fewer opportunities for researchers to make mistakes.*

The extractor builds quantitative capacity through enabling users to:

- Merge different files to collate information into one dataset
- Recode family member relationships - for example, if parents are younger than their children, when it looks like there is a miscode of who is partnered to whom
- Create consistent definitions of variables such as earnings (e.g., coded in weekly terms) and consistent coding frames in cases where these variables change over time
- Merge cross-wave information to promote longitudinal analysis

Having a standard self-curated dataset used across the IFS that contains cleaned, derived variables means findings can be replicated and checked to ensure they are representative and reliable. This increases the credibility of findings, which the IFS prides itself on, through providing an independent and respected site for quality evidence that can inform topical fiscal debates:

> *We want to be a kind of repository of knowledge about policy and how it affects different people outside of the treasury so that we can be a resource there for the public debate without people having to basically accept what they're told by the experts inside government.*

Using *Understanding Society* in this way bolsters capacity and shared knowledge of quantitative methods to answer specific research questions. This enables the IFS to remain at the forefront of advancements in knowledge:

> *Features of the [Understanding Society] data that aren't present in other UK and international datasets, make it useful for us to have impact and to be on the frontier of knowledge.*

Case Study 3: Using the Innovation Panel to foster methodological innovation

Using the Innovation Panel as part of *Understanding Society* provides researchers with the opportunity to test, and be creative with, new methodologies and methods. The Innovation Panel provides a sample of 1,500 households, which enables researchers to test new concepts and to challenge taken-for-granted norms within a discipline.

One user of the Innovation Panel explored a different methodology for measuring and defining successful ageing[17]. Due to the costs and resource associated with conducting a study from scratch, this research would not have been possible without the Innovation Panel. The dominant definition of successful ageing at the time was from a clinical perspective that largely ignored the social viewpoints of the patient experience. To address this, the research team created a vignette study for the Innovation Panel. Respondents of all ages were asked to rank how successful they thought people described in three vignettes were at ageing from 0 (unsuccessful) to 10 (successfully).

Vignettes were created to mirror six commonly used dimensions of ageing (disease, disability, physical and cognitive functioning, and interpersonal and productive social engagement). Participants received three vignettes to assess, describing both men and women. The vignettes were randomly allocated, with the different dimensions of ageing being described in positive or negative terms through the software. Examples of the vignettes used are shown in Figure 8:

[17] This research was conducted primarily by independent researchers, however, there was some involvement from the *Understanding Society* team at the University of Essex, who are named on published outputs.

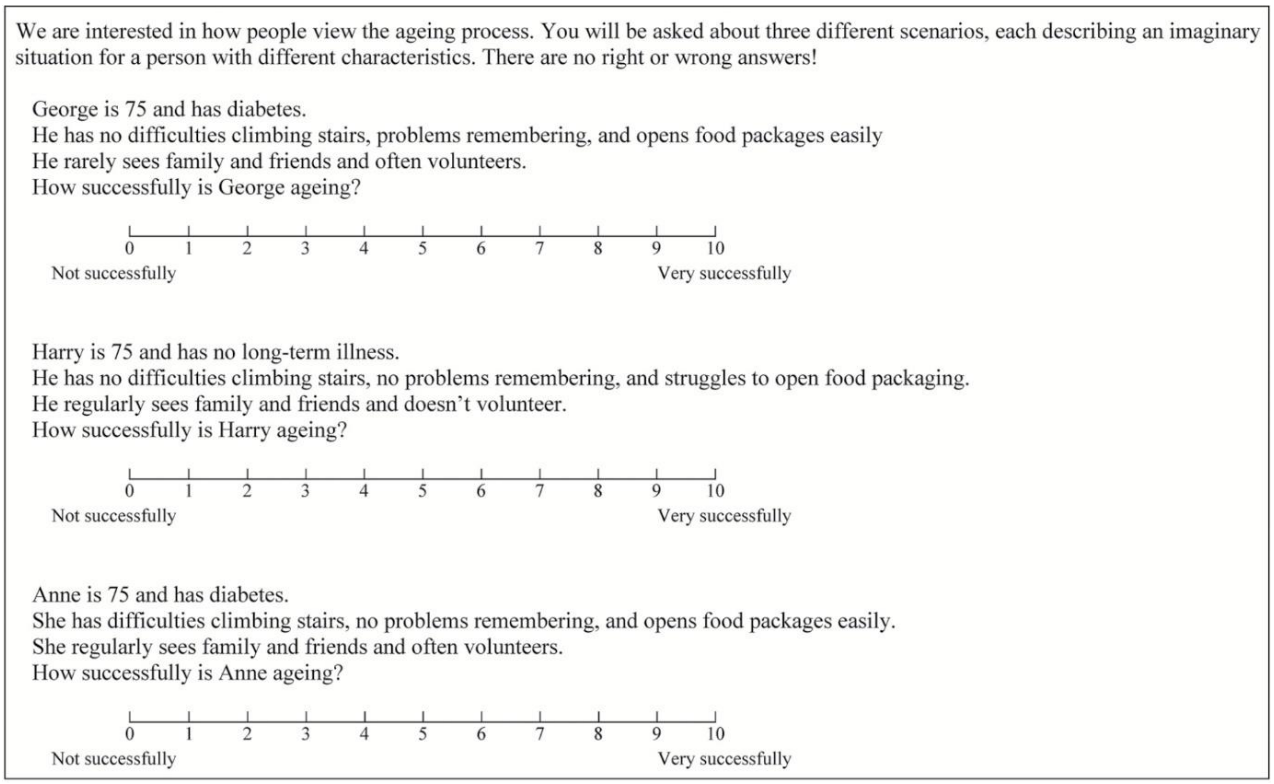


Figure 8: Vignettes for successful ageing

The use of vignettes is an innovative method to study health and ageing. A key benefit of testing new methodologies using the Innovation Panel is to challenge the status quo of evidence and ingrained views within a subject. It also provided the opportunity to derive novel data to bridge the qualitative-quantitative divide. This was achieved by incorporating subjective elements and social perspectives into a large sample of responses. The researchers note that '*the Innovation Panel gives you the freedom to be creative – it is a safe space to do something different*'.

Academic impact was achieved with a widely cited published article. Adopting an innovative perspective on ageing in this study has contributed to a paradigm shift towards understanding the importance that older people place on physical and social aspects of ageing. Moreover, the opportunity to speak with approximately 2,500 people about their perspectives on ageing contributed to patient and public involvement (PPI) initiatives. This has emphasised the need for decisions and outcomes to be geared towards the areas that are significant to the lived experiences of older people, i.e., the social aspects of ageing.

Case Study 4: *Understanding Society* meets the shared needs of research and policy users for research into transport

The unique features of *Understanding Society* enable novel academic, and Government funded projects into transport and travel behaviour. For example, an internationally based academic is carrying out research to understand the link between travel behaviour, migration background and social network dispersion. Social analysts at the Department for Transport are also working on a range of projects including the decarbonisation of the aviation industry and understanding working from home patterns during the COVID-19 pandemic.

The unique features of the data enable users to make connections between different topics:

> *Transport studies is a very interdisciplinary area… people come at it from different disciplines.*

In this sense, *Understanding Society* can promote an interdisciplinary perspective and meet the shared needs of both research and policy users understanding more about travel and transport.

Existing datasets such as the National Travel Survey provide data about travel behaviour and sociodemographic characteristics. However, they do not offer the breadth and richness of data required to fully address specific transport and travel research questions. Collective insights from DfT analysis and academic contributor suggest that *Understanding Society* helps to plug these gaps by offering several benefits:

- It provides a valuable **question set** that enables **detailed information** about air travel, such as how many flights people have taken over the last 12-months and whether friends and family live abroad
- **Longitudinal analysis** can be carried out to understand how people adapt to environmental changes over time and to begin to explore issues of causality
- Travel behaviours can be **linked** to other sociodemographic variables including life events, changes to employment and marital status, private car usage, income, migration status, as well as attitudes towards climate change and questions about wellbeing. The links between environmental and transport topics and attitudinal perspectives can offer a greater understanding of what policies or events mean for different people in different places.
- The **rotating module structure** over a series of data collection waves allows a wider variety of variables to be linked and provides more possibilities for analysis,

including linking to special license files for lower-layer geographical data (LSOA[18]) and available government data on distance to airports

- **Timeliness** of the COVID dataset promotes innovative research on travel behaviour to respond to emerging policy areas. For example, analysis compared the wave 10 mainstage survey data as a benchmark and compared this to COVID wave 3 survey that included questions on working from home and the frequency of travel during the COVID-19 pandemic

- **Large sample size** provides analytical assurance that the data is of high quality and credible. For the Department for Transport, being confident in the quality of the data helps create a narrative regarding the credibility of analysis when attempting to influence policy makers:

> *We're telling our ministers what we've done is robust and we believe in this analysis. –* ***Department for Transport analyst***

Using *Understanding Society* has led to two publications for the academic researcher - one resulted in media coverage about migrants who care about climate change but are required to fly frequently; another contributed to academic theory impact and knowledge by exploring potential causal mechanisms underlying higher levels of long-distance travel activity among residents of urban areas. There are plans for a third publication with a policy angle towards the frequent flyer levy. For this academic, engaging with *Understanding Society* created new opportunities to explore the topic of travel and transport from a novel interdisciplinary perspective.

Findings from using *Understanding Society* have been used by the Department for Transport to show the value of transport investment for the wider lives of people, which can be appealing to policy makers. At a policy level for the Department, the range of variables available in *Understanding Society* enables '*the ability to make connections between different topics*', thus promoting an interdisciplinary approach. Opportunities to carry out longitudinal analysis means that the evidence base can continue to be built to identify trends in travel and transport behaviour. Adopting longitudinal and multi-variable analysis provides valuable and timely insights for the Department to inform changes to policy and practice. The Department plans to exploit the data further in future projects such as looking at the impact of ageing on how people travel and why people 'get greener.'

[18] Lower-layer Super Output Areas – small areas designed to be of a similar population size, with an average of approximately 1,500 residents (650 households). There are 32,844 LSOA in England.

Case Study 5: The validation and testing of a forecasting framework using *Understanding Society* to develop innovative methods and advance statistical methodologies

The sample size, range of variables and longitudinal characteristics of *Understanding Society* means it can answer a range of research questions for a divergent set of organisations. This case study is based on the use of *Understanding Society* data to support research by the Institute of Defense Analyses in Washington, DC – a think tank that is federally funded and provides research for the Department of Defense and other federal departments.

The Institute has developed their own forecasting models to inform military force planning. These models were constructed using the Institute's own data on military personnel. Adopting good practice within forecasting methodology the researchers' used *Understanding Society* as an external data set to validate both their statistical framework and the software tool that implements that framework. Using *Understanding Society* in this way provides an alternative use of the dataset, offering a testbed to improve forecasting models.

For this project, the research team applied their statistical modelling framework to all waves of the BHPS and *Understanding Society*. For these researchers, '*longitudinal survey data such as the UKHLS, with thousands of features and thousands of participants each year, allows us to use this framework to learn about future populations.*'

The Institute of Defense Analyses has their own longitudinal data on members of the Department of Defense and use this to analyse the circumstances of personnel in the US military over time. Researchers developed a tool called FIFE, the Finite-Interval Forecasting Engine. This tool is used with the Institute's own data to inform military force planning by forecasting outcomes such as: whether a service member will separate from the military in each future month; what rank each service member will be each month into the future, given that they are still serving; and the type of discharge (e.g., honourable, medical, bad conduct). In order to validate FIFE and test its capabilities, the Institute needed a longitudinal dataset that could mirror their own military data.

Using the variables that are available in each wave of *Understanding Society*, the analysts chose marital status and urban/rural residence because these variables clearly show change over time. The analysis involved applying a machine learning algorithm (called gradient-boosted decision trees) to forecast outcomes such as when a respondent will be alive in each future survey wave, and, if so, what their marital and urban/rural residence statuses will be. The analysts forecast these outcomes up to four waves ahead. In a separate exercise, the analysts excluded the last four waves from modelling, then used those last four waves to assess their models for accuracy. The results demonstrated that their models could produce high performing individual-level forecasts. The use of an external dataset also meant that researchers were also able to find and fix bugs in the code by applying

FIFE to *Understanding Society*. ln addition this has helped analysts '*explain FIFE and the methods behind it to potential users, and to inspire those potential users.'*

The researchers' used *Understanding Society* to validate their statistical framework and the software tool that implements that framework. Machine learning models produced by that tool are direct evidence of the use of *Understanding Society* to develop innovative methods and encourage advanced statistical methodologies. Rather than the content of the study being important, for this research, it was the structure and size of the dataset that allowed demonstration of the forecasting tool to be possible. This shows a wider strength of the dataset above more traditional uses and forms of analysis. The showcase using *Understanding Society* has also led to a plan for the Institute of Defense Analyses to train new models to inform a controlled trial of service member retention interventions. Researchers are confident future applications of FIFE can produce models… *'as diverse and innumerable as the outcomes observable in current and future longitudinal datasets.'*

Case Study 6: Using *Understanding Society's* biological data to explore stress

The availability of social and biological data within *Understanding Society* has the potential to inform internationally leading interdisciplinary research. Biomarker data was collected in Waves 2 and 3 of the main survey via face-to-face nurse visits. This provides a diverse and rich collection of biological health markers including: measures of blood pressure, weight, height, waist measurement, body fat, grip strength and lung function. Blood samples taken also provide information on fat levels, diabetes markers, immune system inflammation, anaemia measure, liver and kidney function and hormone levels.

A team of researchers have used this biological data to explore an individual's allostatic load (cumulative stress levels) with perceptions of their neighbourhood including crime and safety, a novel area of research in this field[19]. The study sought to explore whether perceived or actual neighbourhood crime and safety have an impact on resident's stress and how this compares to (actual and perceived) deprivation and social capital. A further element explored whether neighbourhood inequalities in stress can be ameliorated by individual education, occupation or income. To explore allostatic load, the team used a range of biological markers available in *Understanding Society* including creatinine clearance rate; Dehydroepiandosterone Sulfate; Insulin-like growth factor; C-reactive protein; Total-HDL cholesterol ratio; Triglycerides; Haemoglobin A1c; Systolic/Diastolic blood pressure; Waist-height ratio and Fibrinogen. The key findings were that inequalities in allostatic load, an objective measure of cumulative stress on the body, are observed in individuals regardless of individual socio-economic position and neighbourhood characteristics. However, the research also suggested that having a higher socioeconomic position may protect against the negative impacts of poor neighbourhood perceptions and conditions.

The above research questions could not have been explored without the presence of the biomarker data in *Understanding Society* – '*looking at allostatic load and pairing that to neighbourhood data…that's been a great thing to do. I'm not sure we could have done that on another dataset*'. Previous research in allostatic stress has focused on subjective self-reported stress scales. The availability of biomarker data linked to social variables and objective area characteristics provided a unique opportunity to '*explore associations using an objective measure of stress which also provides an early warning of later poor health*'.

Incorporating biological data, together with self-reported survey data enables researchers to undertake novel interdisciplinary research to provide robust and empirical evidence - '*the unusual combination of objective biomarker data, questionnaire responses and linked crime and census data'* has not been considered before. The resulting academic

[19] This research was conducted primarily by independent researchers, however, there was some involvement from the *Understanding Society* team at the University of Essex, who are named on published outputs.

journal article is in the process of being published, and it is anticipated that the findings will have an impact across different policy and economic areas including neighbourhood planning and health.

Case Study 7: Using biomarker data to gain new interdisciplinary insights into the interaction between testosterone and social outcomes

One distinguishing feature of *Understanding Society* is the subsample that includes nurse visit and biomarker data. Linking individuals' biological data to demographic and attitudinal variables enables researchers to gain new insights into the interaction between biological and social factors.

One research team wanted to explore the relationship between testosterone and social outcomes. They sought to understand whether there are significant correlations between testosterone levels in men and changes to health and social indicators such as cardiovascular health, diabetes and marital status.

Eight propositions were tested in the data: 1) male testosterone has a diurnal rhythm, 2) low testosterone is associated with obesity, 3) testosterone levels decline after the age of 30 years, 4) married men have lower testosterone than single men, 5) cigarette smoking raises testosterone, 6) testosterone is associated with risk-taking behaviours, 7) low testosterone is associated with likelihood of Type 2 diabetes, and 8) low testosterone is associated with metabolic syndrome.

Understanding Society includes testosterone measurements for 5,940 men gained from nurse visit data collection points in Wave 2/3 (for the General Population Sample and BHPS sample respectively). The measurement of testosterone was crucial to enable this team to address their research questions:

> *Testosterone is not available in many datasets…and as I looked over the data I saw that it would be very useful in evaluating some of these floating propositions in the literature [on testosterone].*

During the analysis, the research team looked at levels of testosterone by age, and how various health and lifestyle outcomes changed for different age groups. This included marriage status, obesity, smoking, risk taking and diabetes. The combination of biological and social data means that research can increase knowledge and understanding which has the potential to inform policy and practice for public health.

Using the combination of biomarker and social data in this way has enabled the research team to recently publish their study. *Understanding Society* data led to an increase academic knowledge. One author said, "we couldn't have done the study without the data. It's that simple".

The findings supported seven of the eight propositions about the characteristics of testosterone in men – no association was found between testosterone and risk-taking behaviours. Collectively, the findings could inform changes to medical practice and help build a body of evidence. For example, testosterone measurements are traditionally taken

in the morning in the belief that levels decline consistently during the day. This research challenges this view as it shows that as men age, daily declines in testosterone levels are less pronounced than previously thought. This has a clinical implication that greater flexibility is possible with timetabling testosterone measurements for older men (aged 60 or older).

Measuring testosterone is infrequently carried out in large social surveys. The methods used by *Understanding Society* to collect the testosterone samples are reliable and well-respected. The research team's findings simultaneously help to validate these data collection methods because so many of the research propositions can be substantiated from the data.

Findings from this research suggest a need for more causal analysis to explore testosterone treatment and potential behavioural consequences of modified testosterone levels. Currently, testosterone measurements are only taken at one time point for the *Understanding Society* data. This means that exploring any longitudinal or causal relationships is not possible. Repeated testosterone measurements would allow consideration of changes in testosterone and how they relate to changes in other variables. This would enable analysis to control for unobserved effects that can be assumed not to change over time, thus bringing the analysis closer to being able to make causal explanations. For example, repeated testosterone measurements would help to generate a greater understanding of the relationship between testosterone and weight and/or marital status over time.

Case Study 8: Interdisciplinarity: challenging the definition of Public Health research

The unique features of the *Understanding Society* dataset promote creative thinking to address major public health issues from an interdisciplinary perspective. The rich variables within the dataset enable public health researchers to consider the diversity of the UK population. Interdisciplinary public health projects have explored social disadvantage and ethnicity, the interrelation of multiple risk behaviours (smoking, harmful alcohol intake, low fruit and vegetable intake and low physical activity), and the social patterning of active travel (walking and cycling) in rural and urban areas. Public health research is inherently interdisciplinary, but for these projects, *Understanding Society* allowed the researcher to:

> *Take account of the dynamic quality of people's lives and the wider diversity of the UK population. So, it was interdisciplinary in that we were looking at how people's circumstances and identities were related to behaviours with profound impacts on people's health.*

Furthermore, *Understanding Society* tracks changes in people's lives, health behaviours and health status over time. Analysing these interconnected elements promotes an interdisciplinary lens because people's health is shaped by their circumstances across their lives, and these variables are clearly measured in *Understanding Society*.

One project used *Understanding Society* data to encourage a change in focus for public health policy - to focus less on separate unhealthy behaviours, like smoking or binge drinking, and more on unhealthy lifestyles where multiple unhealthy behaviours are combined (e.g., smoking, a poor diet, binge drinking and physical inactivity). *Understanding Society* offers the advantage that the four risk behaviours are measured in ways that can be aligned with official health definitions (e.g., 5-a-day targets for fruit and vegetable intake). The dataset also enables researchers to consider the wider social patterns of these behaviours across different axes – by household income and composition, gender, age, education, ethnicity:

> *A lot of public health research hasn't been able to take ethnicity seriously and combine a focus on ethnicity with other dimensions of diversity and inequality. So [Understanding Society] allowed us to widen the frontiers of public health away from a single dimension focus to a more social science focus.*

Using Latent Class Analysis (LCA) methods focused on households with children, the research team showed that the four behaviours clustered together in distinctive ways. This is a new finding, not previously reported. For example, the largest behavioural cluster were non-smoking binge drinking parents, who were physically active; this group were more likely to be socially advantaged (white, married, with higher education and higher incomes). The researcher suggested that this type of analysis was only possible using

Understanding Society due to the granularity of key variables (e.g., distinguishing between smoker categories and frequency of smoking), and the size and richness of data about people's social circumstances, and the large size and household structure of the dataset. Collectively this offered '*a much richer picture of how health behaviours are embedded in people's lives*'.

These projects have had numerous impacts at the policy-making level, particularly in relation to reframing how policy considers health. Some of the above mentioned projects were part of a Department for Health funded research programme, and resulting impact included enabling policy makers to think more about the communities and groups of people who engage in multiple risk behaviours, and to not simply focus on the behaviours themselves: '*the availability of [Understanding Society] facilitates this thinking*.' The researcher emphasised that these projects supported the government and policy community of the time in thinking about people and their lifestyles in a more holistic way, to the point where thinking about people's lives and lifestyles together became more widely understood.

Case Study 9: *Understanding Society* contributes to the debate about reducing the working week

Academic research findings using *Understanding Society* can be the catalyst for sparking important debates on future policy areas, as well as contributing to existing policy discussions. Sociologists have used *Understanding Society* to demonstrate links between changes in working hours policies and the potential positive impact on mental health and wellbeing. The importance of addressing this issue has been emphasised by the continuing COVID-19 pandemic.

The aim of this research was to establish the number of working hours that can improve a person's well-being and achieve positive mental health. All waves of *Understanding Society* and BHPS were used to analyse a sample of over 70,000 people. Changes in work patterns and mental wellbeing were analysed over time. The resulting academic paper generated significant worldwide interest from over 350 media outlets, including all major national UK newspapers such as *The Guardian* and *The Independent* as well as a range of other news and social media sites. This has led to further research being conducted by the team, using the COVID-19 dataset to investigate the impact of the pandemic on the *"mental health of people who were on furlough or reduced hours or lost their jobs."*

There is significant policy debate in the UK around unemployment, working hours and universal basic income. This work using *Understanding Society* has fed into this debate and researchers have made several policy recommendations including reducing the standard working week and distributing the surplus hours to those who need more work. *Understanding Society* was crucial in enabling the researchers to generate policy and academic impact while investigating specific questions to establish whether relationships occur between employment and mental health:

> *Understanding Society allows researchers to drill down into the relationship [between employment and mental health], to reduce all alternative explanations as far as possible. That I would say is academic impact. It provides academic and policy debates with more rigorous evidence.*

There has also been significant economic impact from this work:

> *The research has fuelled more of a discussion about a shorter working week in Britain. As a result, now, there is a large experiment going on in Britain, where many companies are introducing a shorter working week.*

The unique features of *Understanding Society* supported the generation of explanations that are closer to causality, due to the large number of variables and long timeseries, features that are often not available in other datasets. Using *Understanding Society* to

produce evidence that considers the interrelation of variables, whilst controlling for other factors, means the evidence is more robust and can have more credibility in policy debates.

Drawing on *Understanding Society*, has collectively enabled these research projects to identify the most 'effective dose' of paid work needed to positively affect mental health and well-being. Controlling for income, age, and longstanding illness, the main findings across the research projects are:

- There is scope for the working week to be radically reduced
- Reduced hours of work are preferable to job losses
- Working between 1–8 hours a week generates significant mental health and well-being benefits for those who have been previously unemployed or economically inactive.

The research team note that as the pandemic hit, regular Westminster briefing procedures were cancelled and alternative methods were used to disseminate important findings to journalists and politicians. The relevant findings of these research projects were communicated through videos, webinars and the CIPD (Chartered Institute of Personnel and Development) – who could direct the key findings to meetings in the Treasury. The findings from these research projects are policy-relevant to recent events regarding employment and the government's job retention scheme, and such findings have fed directly into the decisions the UK Government made in September 2020 to extend the duration of the furlough system whilst making it more flexible, specifically including allowing working time reductions to protect jobs. For individuals, being able to keep their jobs during COVID-19 over being made redundant, the net effect on the mental health of millions of UK workers was significant and demonstrates the impact that research using *Understanding Society* can have across society.

Case Study 10: Using *Understanding Society* to inform decision-making about economic stability

The Bank of England use *Understanding Society* data to increase understanding of, and to influence economic activity. Analysts use micro datasets (including *Understanding Society*) to explore consumption, spending and cyclical fluctuations in individual/household finances to provide evidence to inform economic stability decision making.

Alongside *Understanding Society*, other datasets, such as the Labour Force Survey, the Wealth & Assets Survey and the Living Costs & Food Survey, are used by the Bank of England to provide a holistic economic perspective across the population. Bank of England analysts emphasised the longitudinal nature of *Understanding Society* as being beneficial to plot the distribution of savings across households with different levels of income and how this changes over time. For instance, a recent blog post outlines the use of *Understanding Society* to understand household's build-up of savings and their marginal propensity to consume (MPC) during the pandemic.

Quarterly findings derived from analysis of *Understanding Society* are presented to the bank's main decision-making committees alongside published bulletin articles. These feed into decision making about interest rates and other important financial stability decisions. *Understanding Society offers* unique opportunities to draw on data to develop distributional parameters in models:

> *Increasingly, we're being asked more distributional questions. To parametrise the models we build, we need to be able to put in distributional parameters, and Understanding Society is going to be one of the few datasets that can provide that. There's just no other dataset in the UK that does distributional savings rates.*

Time-sensitive fluctuations are important to accurately understanding financial activity. Analysts at the Bank of England emphasise that *Understanding Society* could be more widely used if it was more responsive to time-sensitive fluctuations in economic activity, for instance when people's levels of spending and savings change across the year. This would enable more distributional and complex research questions that are integral to the nation's economic agenda to be explored.

The wide variety of variables in the main dataset and the COVID-19 waves also enables the Bank of England to answer specific research questions that seek to address new social problems. One example is the link between the furlough scheme and working from home, and productivity during COVID-19:

> *Understanding Society is really instrumental in helping our thinking around how remote working is likely to affect productivity*

Case Study 11: Using *Understanding Society* to inform and develop the policy evidence base

Policy departments within government and public sector organisations strive to use robust, quality assured data that can inform future policy. Pathways to policy impact are often complex, and tracking precise impact is difficult. The breadth and range of variables covered in *Understanding Society* provides assurances to policy makers about the quality of data:

> *'I don't think there's anything, anywhere, comparable that has the same sample size, the whole range of income variables that we need' (DWP).*

This case study documents the experiences of individuals within two government organisations – the Department for Work and Pensions (DWP) and the Cabinet Office – to showcase the impact their work has had on informing new policy and practice.

Understanding Society is used by the DWP to bridge the gap in information about types of people in the Lifetime Labour Market Database (L2) and the Family Resources Survey. By identifying an individual in administrative data on pensions contributions who is similar to a person in the Family Resources Survey, *Understanding Society* can be used to match these two people and provide a more detailed understanding of their characteristics and financial experiences. This is important so that the DWP can obtain full information about a 'type' of person and use this information for supporting recommendations for policies relating to pension contributions.

Recent DWP research explored the impact of workless households on children, together with the characteristics of workless individuals. This would not have been possible without using *Understanding Society* data because existing administrative datasets cannot be linked to other variables of interest such as health and employment.

For analysts working in the Cabinet Office, *Understanding Society* helps provide the evidence required to endorse suggested policy directions and priorities that should be taken forward at a ministerial level. The large-scale, longitudinal sample provided in *Understanding Society* enables analysts to treat findings with more confidence. For example, the *Understanding Society* COVID-19 dataset has been used alongside analysis from Office for National Statistics (ONS), Public Health England (PHE) and academics to analyse the link between ethnicity and instances of COVID-19 and the implications for policy - examples include an analysis of vaccine sentiment and the COVID health inequalities experienced during the pandemic.

A key role of *Understanding Society* in helping to achieve policy impact is to support the development of the evidence base. Maintaining large sample sizes is key to supporting this activity. However, it was noted by the DWP that the sample size for non-working households is diminishing, and the dropout is not reflective of a random selection of non-working families.

The strength of using *Understanding Society* to form the evidence base rests on its longitudinal structure and large sample size, which enables policy questions to be explored in more detail through group-by-group analysis. Whilst the workings of government are complex and tracking the policy impact can be difficult, both organisations emphasised that having timely research, robust data and telling a convincing story around people's experiences can be powerful for policy makers – all of which are made possible through using *Understanding Society*.

Case Study 12: Third Sector organisation uses *Understanding Society* to campaign for policy change

Age UK use *Understanding Society* data to inform both fundraising and policy work. The datasets enable them to explore key topics associated with older people, including health, care, poverty, employment and loneliness. Compared to other social surveys that specifically include older adult samples (e.g., English Longitudinal Survey of Aging), *Understanding Society* has a larger sample of adults aged over 65, but has fewer questions specific to the needs and/or interests of older people. Whilst there are some measures in *Understanding Society* that have less detail than other surveys, what makes *Understanding Society* powerful for addressing the research needs of a charity such as Age UK is the ability to look at different dimensions of people's lives together:

> *The great benefit of a survey like Understanding Society, is, you might not have such detail on housing, but I can look at different dimensions of people's lives together…[such as] the relationship between their housing and their mental health.*

Understanding Society supports different types of analysis that Age UK carry out to develop fundraising campaigns and inform policy work. For example, fundraising campaigns need to provide headline statistics about how many people are in a particular situation (e.g., how many older people are lonely, how many people are carers). The large, nationally representative sample provided in the *Understanding Society* dataset is ideal for this purpose. A greater level of granularity is required for research intended to inform policy, and here the wide range of variables available in *Understanding Society* is advantageous. Age UK have created an Index of Wellbeing in Later Life using *Understanding Society* which provides policy makers and the general public with information about what is important for older people's wellbeing. *Understanding Society* was used to support this index because it was the available data source which covered the largest numbers of factors that affect wellbeing. Age UK have also used *Understanding Society* to identify characteristics or circumstances associated with loneliness which enables a more detailed understanding about people at greatest risk and the associated causes. Drawing on data from different waves of *Understanding Society* allows analysis to consider the impact of life events, for example, exploring how bereavement affects loneliness by comparing outcomes in the waves before and after it occurs.

Data provided in *Understanding Society* supports Age UK to provide credible headline statistics for press releases:

> *[Understanding Society is] well enough known in policy maker circles that they accept the authority of statistics generated using the data…they're not official statistics but Understanding Society is recognised as a reliable source of data'.*

Tracking the impact of statistics once in the public domain is difficult, but they can help to change the narrative and increase awareness about specific living conditions and situations for older people.

Understanding Society can also help to foster collaborative activities with other organisations, such as Carers UK; where resource is shared across organisations with similar interests and aims. Advanced statistical skills are required to proficiently use *Understanding Society*. These can be hard to source, and third sector organisations may struggle for capacity. Collaboration between organisations can enable skills to be shared and statistics produced that benefit their campaigns and policy interests.

The ability of charities to influence policy can, at times, be hampered by Government timescales. This in turn can hinder more nuanced analysis:

> *Often, we're very reactive. An announcement will be made and we need to work quickly to campaign to influence the more detailed policy making. So we don't have time to do detailed analysis, but we need enough analysis to provide evidence to support our arguments, and to start making noise. We tend to need simple-to-understand statistics rather than subtlety.*

Drawing on data from *Understanding Society*, third sector organisations, such as Age UK, can illustrate key headline statistics to campaigners, Members of Parliament, civil servants, and the public- (via the media). These statistics can push forward policy conversations to ensure that important societal issues are at the forefront of national and policy-maker conversations.