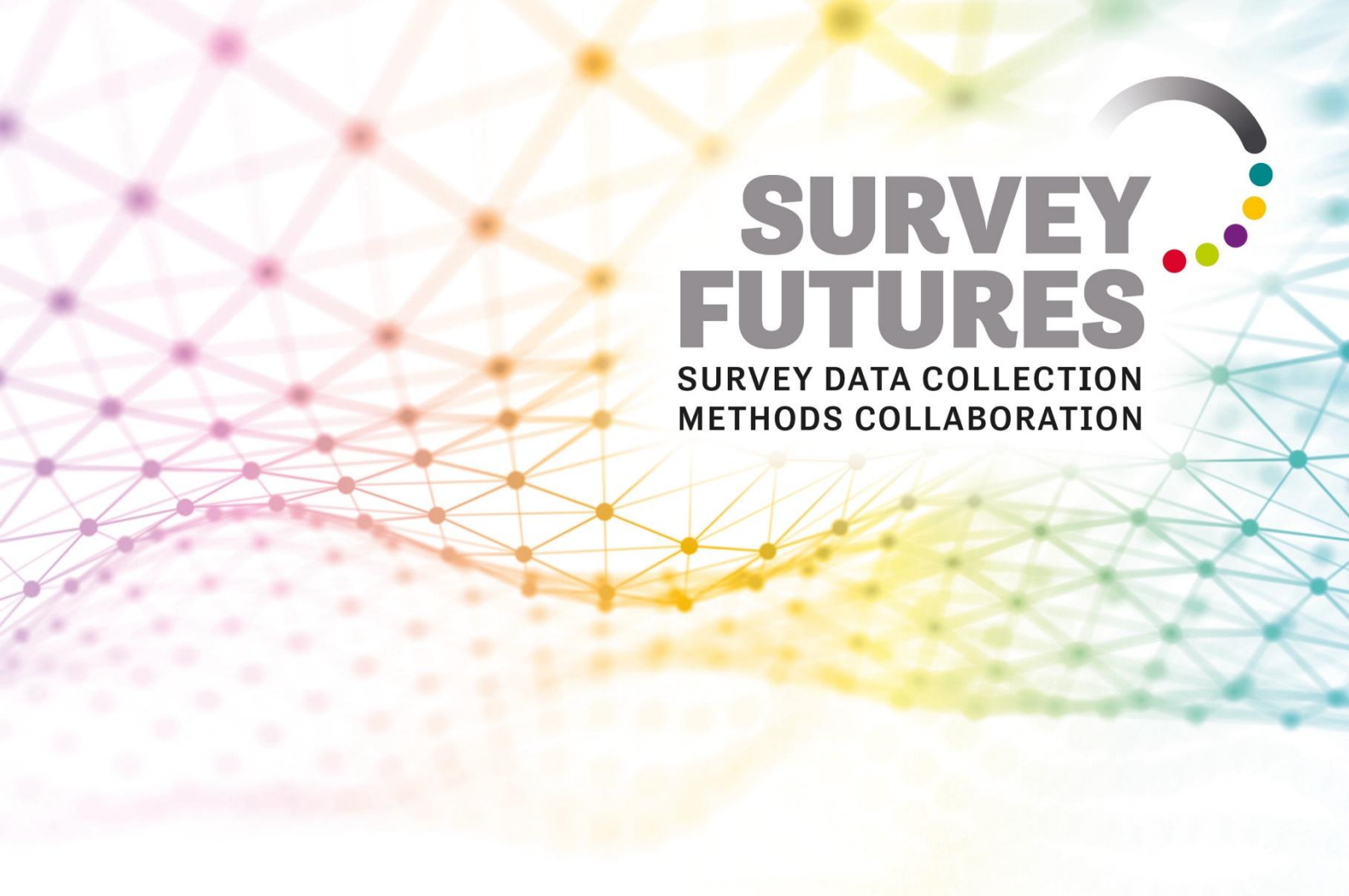




SURVEY FUTURES

**SURVEY DATA COLLECTION
METHODS COLLABORATION**

Report 17: How the RD MF could potentially be used to improve the design and cost-effectiveness of the TLFS: A Q&A

The Office for National Statistics

July 2026

www.surveyfutures.net

Contents

1. Introduction	3
2. Background	4
3. Purpose and scope	5
4. Where could the RDMF add value to the TLFS?	6
5. Risks, limitations and constraints	8
6. What would be required in practice?	9
7. Next steps	10

Survey Futures is an Economic and Social Research Council (ESRC)-funded initiative (grant ES/X014150/1) aimed at bringing about a step change in survey research to ensure that high quality social survey research can continue in the UK. The initiative brings together social survey researchers, methodologists, commissioners, and other stakeholders from across academia, government, private and not-for-profit sectors. Activities include an extensive programme of research, a training and capacity-building (TCB) stream, and dissemination and promotion of good practice. The research programme aims to assess the quality implications of the most important design choices relevant to future UK surveys, with a focus on inclusivity and representativeness, while the TCB stream aims to provide understanding of capacity and skills needs in the survey sector (both interviewers and research professionals), to identify promising ways to improve both, and to take steps towards making those improvements. Survey Futures is directed by Professor Peter Lynn, University of Essex, and is a collaboration of twelve organisations, benefiting from additional support from the Office for National Statistics and the ESRC National Centre for Research Methods. Further information can be found at www.surveyfutures.net.

Research Strand 1 of Survey Futures (“Enhanced sampling frames and procedures”) focuses on the challenges associated with our dependency on using the Postcode Address File as a sampling frame for random probability surveys. Within this strand, the Office for National Statistics explores the potential of using its Reference Data Management Framework to improve the design and cost-effectiveness of the Transformed Labour Force Survey.

The suggested form of citation for this report is: The Office for National Statistics (2026) ‘How the RDMF could potentially be used to improve the design and cost-effectiveness of the TLFS: A Q&A’, Survey Futures Report no. 17. Colchester, UK: University of Essex. Available at <https://surveyfutures.net/reports/>.

1. Introduction

This report is presented in a question-and-answer (Q&A) format to provide a clear and accessible overview of how the Office for National Statistics' Reference Data Management Framework (RDMF) could potentially be used to improve the design and cost-effectiveness of the Transformed Labour Force Survey (TLFS). The questions are grouped thematically, beginning with background information on the TLFS and the RDMF, before considering the purpose and scope of the work, where the RDMF could add value, the main risks and constraints, practical requirements, and possible next steps.

2. Background

Q: Could you provide an overview of the Transformed Labour Force Survey?

A: The Transformed Labour Force Survey (TLFS) is the UK's next-generation labour market survey, designed to replace the current Labour Force Survey (LFS). It is an online first, more flexible survey that aims to address declining response rates, improve data quality, and deliver more timely and granular insights into the labour market.

Further information about the TLFS can be found on the ONS website at:

<https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/methodologies/transformedlabourforcesurveyuserguidance>

Q: Could you explain what the RDMF is?

A: The RDMF (Reference Data Management Framework) is a secure ONS infrastructure designed to link administrative data from across government in a consistent and reusable way. It produces reference indexes for people, businesses and locations to support statistical production and research. A critical feature of the RDMF is that it de-identifies data: replacing names, addresses and other direct identifiers with reference numbers before analysts access the data. This means individuals cannot be directly identified, and the RDMF operates within strict legal, ethical and governance controls. The RDMF is a linkage-first infrastructure that supports the construction and enhancement of downstream assets. For example, the Business Index underpins the development of the Statistical Business Register, while the Demographic Index supports population statistics.

Further information can be found in the following Survey Futures reports:

Office for National Statistics (2026) 'Reference Data Management Framework project - what is the RDMF?', Survey Futures Report no. 13. Colchester, UK: University of Essex. Available at <https://surveyfutures.net/reports/>.

Perera Z M, Nicolaas G & Waugh N (2026) 'How could the Reference Data Management Framework improve survey quality?', Survey Futures Report no. 15. Colchester, UK: University of Essex. Available at <https://surveyfutures.net/reports/>.

3. Purpose and scope

Q: Why is the TLFS being considered in relation to use of the RDMF in social surveys?

A: The TLFS provides a strong test case because labour market surveys are under significant pressure from falling response rates and rising costs. We want to consider how the RDMF could improve quality and efficiency in the TLFS, particularly by enhancing how we sample addresses. As one of the largest and most high-profile social surveys, the TLFS also offers a useful reference point for understanding how similar approaches might be applied more widely. The aim is to assess where the RDMF could add value in practice and to identify the operational, methodological, legal and ethical opportunities and constraints that may limit or shape its use.

Q: What is the scope for considering how the RDMF could be used in the TLFS, and which possible uses are outside scope?

A: The TLFS is an address-level survey based on a sample of addresses and would likely remain so, even if the RDMF is used to enhance its design. Moving to a person-level design would have significant implications across the end-to-end survey process - including sample design, weighting methodology and operational delivery of the survey. This Q&A therefore focuses on how the RDMF could enhance the existing address-based design in a way that builds on current TLFS methods, rather than on transitioning to an individual-based approach. It explores the trade-offs involved (for example, potential efficiency gains versus implementation complexity) and the conditions under which RDMF use would be feasible, proportionate and publicly acceptable. The intention is to provide a realistic assessment of where the RDMF could strengthen the TLFS, as well as the safeguards, governance and delivery requirements needed for responsible use.

Q: What are the current challenges facing the TLFS that make this worth exploring?

A: The key challenges include maintaining representativeness as response rates decline, managing rising fieldwork costs, and preserving the coherence of time series during periods of transformation. At the same time, there is increasing demand for more timely and granular labour market statistics. These pressures are compounded by the need to balance innovation in data collection methods and technology with robust quality assurance, alongside the requirement for upfront investment before longer-term efficiencies can be realised. The longitudinal design of the TLFS introduces additional complexity, as retaining the sample across waves is critical to preserving data quality. Attrition over time can introduce bias if certain groups are more likely to drop out, affecting both representativeness and consistency across waves.

4. Where could the RDMF add value to the TLFS?

Q: What kinds of improvements to the TLFS might the RDMF realistically support?

A: Because the TLFS already uses AddressBase for sampling, independently of the RDMF but as part of the Location Index, there are opportunities to link sampled addresses to auxiliary information derived from administrative data, enabling a more dynamic and informed sampling design.

Beyond this, the RDMF could enable better comparison of the characteristics of respondents and non-respondents, informing more effective non-response strategies. In the longitudinal context of the TLFS, this could be extended to understanding attrition across waves – it may be possible to identify systematic differences between those who remain in the survey across each wave and those who drop out over time. In turn, this evidence could inform improved weighting strategies, adjustments and survey design decisions aimed at mitigating attrition effects.

More broadly, address-level linkage to administrative data rather than relying on higher-level aggregate data could support more targeted adaptive survey design, address-level stratification and weighting, as well as improvements in downstream processing and quality assurance. Together, these uses point to the added value of the RDMF's application across multiple stages of the survey process, rather than as a single isolated intervention.

Q: How might use of the RDMF improve the representativeness of the TLFS?

A: Although AddressBase already provides strong coverage, the RDMF could add value by helping to identify ineligible addresses. For example, commercial properties (via linkage with the Business Index) and complex addresses which may benefit from alternative data collection approaches (for example, communal establishments or houses of multiple occupancy).

The RDMF could also support efforts to minimise non-response error. Linked administrative variables could enable adaptive fieldwork strategies, targeting groups with lower response propensities. In addition, adjustment for non-response error could be strengthened by using auxiliary variables that are more strongly correlated with the survey outcomes and are more timely, thereby improving weighting models and reducing residual bias.

Q: How could the RDMF help the TLFS to better measure the labour market?

A: The RDMF could strengthen the integration of survey and administrative data to enhance measurement of the labour market. Administrative sources may provide more timely or more precise information for certain outcomes than survey responses. Combining these sources with the survey data could therefore support richer labour market indicators and models, particularly for domains where survey-only estimates are more challenging. This may be particularly beneficial for variables prone to reporting error or volatility, such as industry, where collection is more complex in an online context. Respondents may have difficulty accurately identifying or describing their employer's industry, leading to inconsistencies or misclassification. These challenges may be more pronounced for certain types of workers, such as those in large or multi-activity organisations, or workers in less clearly defined sectors.

Administrative sources, where industry information is derived from business systems, may therefore offer greater consistency and stability over time for these groups.

There may also be opportunities to replace some survey questions with administrative data. This could reduce respondent burden and minimise measurement error associated with self-reporting. However, identifying variables which could be replaced with sufficient confidence would require a substantial assessment of the quality, coverage and definitional alignment of the administrative data. In practice, the number of variables meeting the required threshold is likely to be limited, meaning that any reduction in survey length would be minimal. The integrated use of survey and administrative data is therefore likely to be a more feasible first step than question replacement.

Q: Could the RDMF also support the processing of TLFS data?

The RDMF has the potential to play a valuable role in the downstream processing of TLFS data. For example, administrative variables could act as reliable reference points for identifying inconsistencies or potential errors in survey responses, allowing for more targeted and efficient editing to improve data quality.

The RDMF could also enhance imputation by providing rich auxiliary variables to inform models used to fill missing or invalid responses. Using administrative data that are closely related to key labour market outcomes, such as earnings, could improve the accuracy and plausibility of imputed values, thereby reducing bias and improving overall data quality. However, the extent of this improvement is likely to vary across different types of workers. Administrative data on earnings is likely to be more robust for employees paid through PAYE, but coverage may be more limited or conceptually different for self-employed individuals or those with multiple jobs.

The RDMF could also support improved adjustment for non-response error. Richer and more timely auxiliary variables, particularly those strongly correlated with labour market outcomes, could be incorporated into weighting models to improve their effectiveness and reduce residual bias in the final estimates.

Q: How might the RDMF improve the overall cost-effectiveness of the TLFS?

A: Use of the RDMF could improve cost-effectiveness by enabling more efficient targeting of fieldwork effort and reducing wasted resource. For example, better identification of ineligible or non-residential addresses would reduce unnecessary contact attempts, while improved classification of complex addresses could support more efficient collection strategies from the outset. Cost savings could also arise through more effective allocation of fieldwork resources. By using linked administrative data to predict response propensities, effort could be concentrated on harder-to-reach groups where additional contact is most valuable, while reducing effort where the likelihood of response is already high. This type of adaptive approach could help improve response rates and achieved sample sizes at lower overall cost.

5. Risks, limitations and constraints

Q: What are the main limitations of using the RDMF in this context?

A: The main limitations reflect that the RDMF uses administrative data sources, which are not designed primarily for statistical purposes. Key labour market variables are not consistently captured, and available variables may be imperfect proxies, limiting their usefulness for analysis or adjustment. There may also be issues of timeliness and stability – administrative sources can be updated at different frequencies, and changes in underlying systems or processes may introduce discontinuities that are difficult to manage within a survey context. As with other administrative data resources, coverage and quality may vary across different parts of the population. Certain groups, such as highly mobile populations or those in more complex housing situations, may be less well represented, which can constrain its effectiveness in improving representativeness.

The quality and completeness of RDMF linkages will also constrain its usefulness. Errors in linkage, lags in administrative data being updated, or incomplete coverage could lead to misclassification of addresses or groups, undermining both efficiency and representativeness. Finally, there are practical and governance considerations on how the RDMF can be used and the need to maintain public trust. These factors may limit the extent to which it can be integrated into core survey processes. However, a key strength of the RDMF is that it provides a structured framework for understanding, assessing and managing many of these issues more systematically, and at significantly less cost, than would otherwise be possible.

Q: Are there any other potential barriers to using the RDMF in the TLFS?

A: The impact on processing timelines is an important consideration. Incorporating the RDMF into processing would introduce additional steps into an already tightly constrained timetable, where turnaround times for data processing and delivery are extremely tight. Any new processing stages involving the RDMF would need to be carefully designed to operate efficiently within these timelines, without delaying outputs. While this is not an insurmountable challenge, it would require careful planning to ensure the benefits of using the RDMF could be realised without compromising the timeliness of the survey's key outputs.

Q: What trade-offs should be recognised when assessing the benefits of using the RDMF?

A: There are several important trade-offs to consider. Introducing differentiated collection strategies may introduce bias if mode or contact approach is correlated with key outcomes such as employment status. Any efficiency or cost gains must therefore be balanced against the risk of introducing additional systematic error. And while adaptive fieldwork can improve efficiency, there are limits to how much response behaviour can be predicted and influenced. Beyond a certain point, additional targeting may yield only marginal gains while increasing operational complexity.

There are also cost considerations. Developing, maintaining and integrating RDMF-based survey processes will require some further investment and ongoing resource, which must be weighed up against the expected efficiency savings. Finally, increased design complexity may make the survey harder to explain to users, potentially affecting trust and the interpretability of outputs.

6. What would be required in practice?

Q: What are the practical steps required to be able to use the RDMF in the design and implementation of the TLFS?

A: We need to build a comprehensive understanding of the quality and use of the administrative data within the RDMF compared to the survey data. This would inform whether, and how, different administrative data sources could be used in the design and implementation of the TLFS. We could then explore opportunities for using the RDMF to facilitate linkage to administrative data to enhance different elements of the end-to-end survey process, through piloting proofs of concept and iterative research alongside existing TLFS methods to understand the constraints and identify further work.

At present, the RDMF already provides a strong linkage backbone with further development needed to support survey-specific applications at scale. Developing coverage assessments, linkage diagnostics and quality insights would be critical, some of which already exist. Robust governance and clear pathways of use of the RDMF need to be established, including time to collate substantial amounts of metadata, quality information and feedback pathways. Time and resource are needed to invest in the appropriate skills, systems and documentation. In addition, survey-specific pipelines using the RDMF would require coordinated effort across methodological, analytical and data engineering teams, and any potential discontinuities arising from changes would need to be carefully managed.

7. Next steps

Q: How do you envision the future of the TLFS with the potential use of the RDMF?

A: The RDMF should not be viewed as a substitute for the TLFS, but as a foundational capability that could enhance and extend it. Surveys will remain essential for capturing information that administrative systems do not observe, as well as for providing the required geographic coverage. The value of the RDMF lies in its ability to help the TLFS to better define its target population, to reach that population more efficiently, and to produce results that are more robust, interpretable and trusted.

Q: Which potential uses appear most realistic in practice?

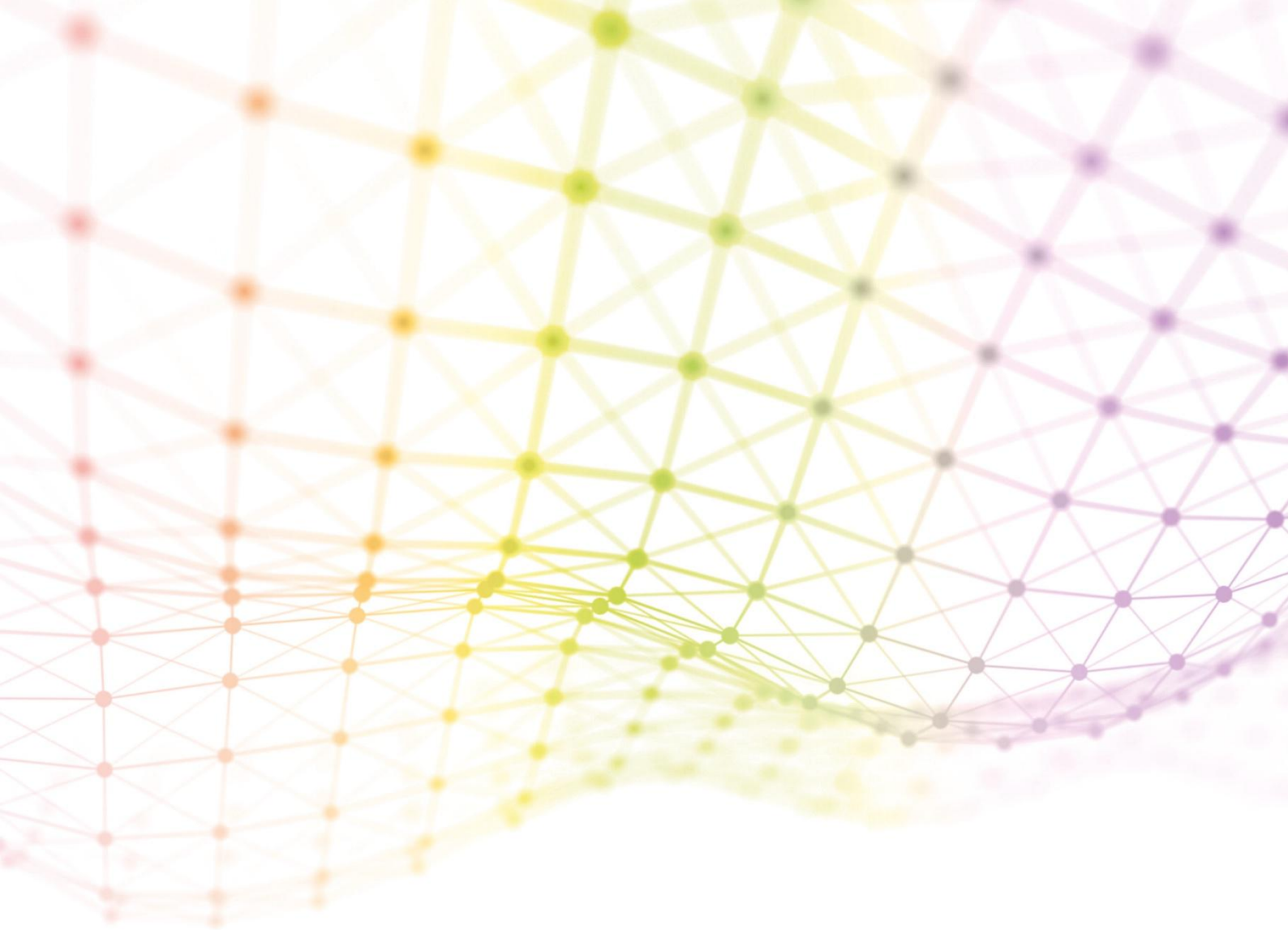
A: Use cases based on address-level linkage are the most feasible currently. These align with the existing TLFS sampling design and operational processes, and therefore avoid the need for high-risk transformations and re-design. In particular, improving the quality of the sampling frame and enhancing adaptive fieldwork strategies are likely to be realistic in the near term. These uses are relatively low risk because they build on established processes and do not require person-level linkage. More ambitious uses would require further development work and careful evaluation before being considered viable.

Q: What should happen next if ONS decided to take this forward?

A: If ONS decides to take this forward, a phased and evidence-led approach would be taken. In the short term, the focus should be on developing and testing priority use cases that align with the current TLFS design. These lower risk applications would provide early evidence on feasibility and impact. Alongside this, investment would be needed in strengthening the underpinning data and infrastructure of the RDMF. Clear governance and documentation will be critical for supporting transparency.

A programme of evaluation would need to be run in parallel, comparing outcomes with and without RDMF-based enhancements. Evidence from these evaluations, including impacts on response rates, bias, precision and operational costs, would then inform decisions about scaling up or refining specific approaches.

Any progress would need to be incremental and proportionate, prioritising low-risk, high-value improvements while building the evidence base needed to support any more ambitious developments over time and contributing to a more integrated, data-driven survey system.



University of Essex



University of
Southampton



Economic
and Social
Research Council

NCRM NATIONAL CENTRE FOR
RESEARCH METHODS

Office for
National Statistics

UCL

WARWICK
THE UNIVERSITY OF WARWICK

MANCHESTER
The University of Manchester

CITY
UNIVERSITY OF LONDON
UNIVERSITY OF LONDON

**National Centre
for Social Research**

LSE LONDON SCHOOL
OF ECONOMICS AND
POLITICAL SCIENCE

Usher University

Unil
UNIL | Université de Lausanne

Ipsos

verian
Formerly Kantar Public

www.surveyfutures.net