# Reducing bias in household survey estimates using enhanced frames

Bruce Fraser[1], Paul Schubert[1] and Anders Holmberg[1]

**Abstract**

Household surveys run by the Australian Bureau of Statistics (ABS) use a sample frame drawn from an Address Register, which as its name implies is simply a list of physical residential addresses without any details about the household or its members. This means essentially surveying 'blind', without knowing about the occupants before contact is made to the household through physical or electronic means. Previous sample designs have though been able to draw on information at an area level, for example from the Census, to target samples towards areas with higher concentrations of particular subpopulations of interest.

As response rates decline over time, and the ABS transitions away from a reliance on face-to-face interviewing to a digital-first approach (eg eforms supplemented by telephone interviews) there is a greater need to understand how well the responding sample represents the populations of interest. In particular, we are interested in predicting and mitigating any non-response biases that may arise, in order to maximise the quality of the survey estimates.

The ABS now hosts the Person Level Integrated Data Asset (PLIDA), a secure database combining mostly administrative information at a person (and household) level on health, education, government payments, income and taxation, employment, and population demographics (including the Census) over time. This opens up new possibilities for minimising non-response biases in ABS household survey estimates.

By augmenting the sample frames with PLIDA (or a subset of it), our sample designs can take advantage of targeting households who are likely to have occupants with characteristics of interest, and we can also model probabilities of response and oversample households with lower response propensities. PLIDA-enhanced survey frames can also be used during survey estimation, to drive follow-up strategies or mode offers eg by using response propensities by mode. A third way of using PLIDA-enhanced survey frames is in estimation, to compare the responding sample to the populations of interest, and identify and adjust for biases in the responding sample.

This paper will discuss these three uses of enhanced survey frames to reduce non-response biases in more detail.

**Keywords**

household survey frames, auxiliary data, non-response bias.

---

[1]Australian Bureau of Statistics, Belconnen, ACT, Australia

**Corresponding author(s):**

Paul Schubert, Australian Bureau of Statistics. Locked Bag 10, Belconnen ACT 2616, Australia. Email: paul.schubert@abs.gov.au

## 1. Introduction

Australia is one of the countries that does not have the benefit of a population register to assist in conducting surveys of people or households.

The best available source of information in Australia from which to draw household survey frames is an Address Register (Australian Bureau of Statistics 2020). The Address Register is itself compiled from a variety of sources, but essentially each record contains only an address, the use of the address (eg residential or business) and type of dwelling structure information. Crucially, it does not contain any information at all about the people who live at the address.

Previous sample designs have been able to draw on information at an area level, for example from the Census, to target samples towards areas with higher concentrations of particular subpopulations of interest. However, when it comes to selecting and surveying individual households, we have no information about the occupants before contact is made to the household (generally through a physical letter in the mail).

Increasing availability of administrative data opens up possibilities for knowing more about the likely occupants of households before sampling, thereby improving the efficiency and effectiveness of household survey estimation. This is important in addressing challenges of declining response across household surveys in Australia and worldwide, and in being able to predict and mitigate any non-response biases that may arise as we transition away from a reliance on face-to-face interviewing to a digital-first approach (eg eforms supplemented by telephone interviews).

In this paper we discuss how administrative data is attach to the Address Register, and how we then use this augmented Address Register for sample design, enumeration and estimation improvement.

## 2. Augmenting with auxiliary data to create Data Improved Frames from the Address Register

Figure 1 below shows an outline of the process to create what we call DIFFAR: the Data Improved Frames from the Address Register. The process is carried out in a secure environment called *StatsLab*, which is an internal clone of the *DataLab* environment used to provide controlled access to unit record data to researchers.

Working our way from right to left in Figure 1, the administrative data is held in a secure asset called the Person Level Integrated Data Asset or PLIDA. This contains person-level information from across Government on health, education, government payments, income and taxation, employment, and population demographics (including the Census) over time (Australian Bureau of Statistics 2023). The PLIDA person-based data contain residential addresses of the people which need to be standardised and geocoded in much the same way as the Address Register so that they can be linked by address.

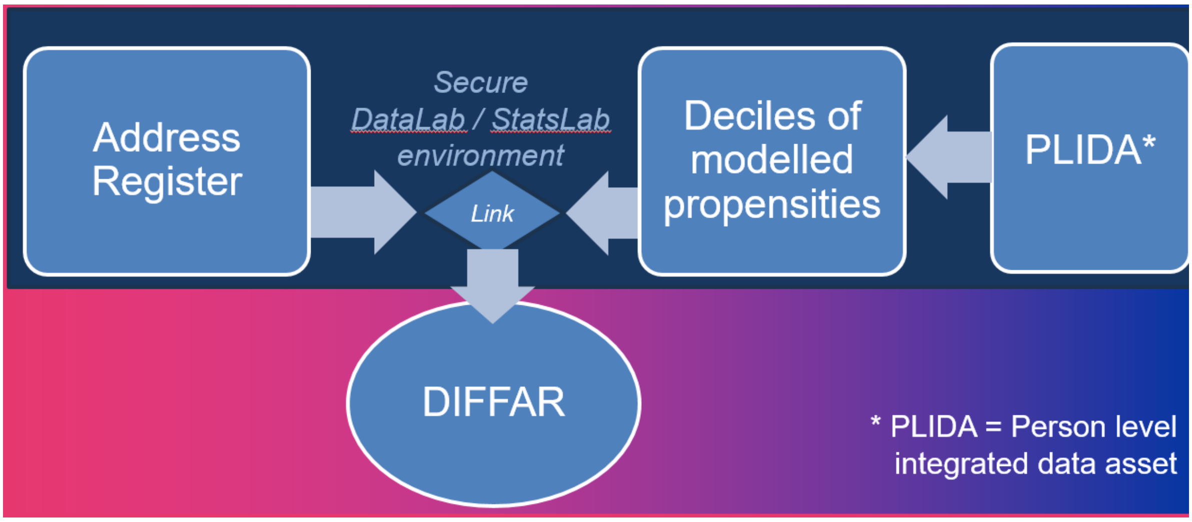


*Figure 1: a schematic representation of the process to create Data Improved Frames from the Address Register (DIFFAR)*

The next step is that various propensities are calculated based on models which have been developed using linked historical survey data and PLIDA. One of the most important propensities is the propensity of a household to respond via an online form, but other propensities relate to various characteristics of the household:

- propensity to receive a government pension or benefit
- propensity that the household includes at least one person with a disability
- propensity that it is a single-adult household
- propensity that it is a couple-only household
- propensity that the household includes at least one child aged 0-14
- propensity to be in the lowest decile when it comes to income of the household
- propensity to be in the lowest decile of a summary indicator of socio-economic status.

The propensities are then converted into deciles which are attached to each person (via the household) in the PLIDA dataset. This is done for privacy purposes, so that no individual PLIDA data leave the secure *StatsLab* environment. There may be some loss in utility in using deciles rather than raw PLIDA data items but this is expected to be small.

The PLIDA with propensities is then linked by address to the Address Register and the resulting Address Register enhanced with the propensities becomes the DIFFAR.

**3. Using DIFFAR for sample selection**

Many of the subpopulations captured by the propensities (eg low income households, households containing a disabled person) are known to have lower response rates to our surveys than the general population, particularly low online (ie self enumerated rather than interviewer assisted) response rates.

We can use DIFFAR in two ways in sample selection:

- if our survey aims to target particular subpopulations, we can give greater probabilities of selection to those households with higher propensities to contain people within those subpopulations
- we can oversample households with lower propensity to respond online so that they are not under-represented in the final responding sample.

In this latter case, the sample selection weights are adjusted by a factor which is the ratio of the target fully responding sample in stratum *h* to the expected number of responses to be achieved in stratum *h* (estimated from the modelled response propensities) ie the initial selection weights become

$$w_{i,h} = \frac{n_h^t}{n_h^r} \times \frac{N}{n^t}$$

where

$n_h^t$ is the number of fully responding persons (FRP) target from stratum *h*

$n_h^r$ is the number of FRP achieved from stratum *h*

*N* is the Estimated Resident Population of in-scope persons

$n^t$ is the FRP target across all strata.

**4. Using DIFFAR for sample enumeration**

Historically, the only dimension of response rates that have been monitored for ABS Household Surveys is geographic eg State/Territory, or metropolitan / ex-metropolitan. System limitations have made it hard to see and monitor the data in real time, only in retrospect which is too late for use in enumeration.

Having the PLIDA data matched to the survey frame and sample helps us to monitor response by characteristics of household and people in it – this enables us to prioritise non-response follow-up towards groups where response is lagging. We have recently implemented this in our household survey enumeration monitoring.

In the future we could go further and use an adaptive design approach by changing our strategy for enumeration in advance based on household characteristics and likelihood of enumeration success. For instance, if we know a household is unlikely to respond online from the outset, we may go straight to a different enumeration mode such as a phone call or visit for this household. Alternatively, we could change the tone in our reminder correspondence accordingly – for example, push harder and more firmly for responses from households we know are more likely to be able to respond via an online form.

**5. Using DIFFAR for estimation**

Both the bias and variance of survey estimates can increase if we have unequal response rates over different portions of the sample

Previously, we were very limited in the benchmarks we could use to adjust for representation of the population in the responding sample – typically, just demographics were available. Some data items of interest would not necessarily be strongly correlated with these benchmarks, so if response rates dropped we would be exposed to risk of bias in survey estimates.

PLIDA can provide a much richer set of benchmarks to adjust for under-representation of various segments of the population eg low income or socio-economic households; government pension or benefit recipients.

## 6. Results

### *6.1 Sample selection*

The DIFFAR approach was used to boost sample selections in lower-response deciles in the 2024 (Australian) Time Use Survey. An aim of the design was to achieve approximately 1000 diary days in each of the response deciles (ranging from low propensity to respond deciles, to high propensity to respond).

| Response decile | Target number of diary days | Using boosting strategy | | Without boosting strategy | | Relative weights | |
|---|---|---|---|---|---|---|---|
| | | % of diary day target achieved | Total diary days collected | % of diary day target achieved | Total diary days collected | Weight with boosting | Weight without boosting |
| **All / Total** | **10,038** | **179%** | **18,010** | **177%** | **17,809** | **1.00** | **1.00** |
| **1** | 1,009 | 222% | 2,244 | 116% | 1,174 | 0.81 | 1.52 |
| **2** | 1,008 | 201% | 2,028 | 138% | 1,393 | 0.89 | 1.28 |
| **3** | 989 | 178% | 1,763 | 145% | 1,434 | 1.01 | 1.22 |
| **4** | 995 | 177% | 1,765 | 163% | 1,624 | 1.01 | 1.09 |
| **5** | 1,003 | 141% | 1,414 | 139% | 1,395 | 1.27 | 1.28 |
| **6** | 988 | 133% | 1,314 | 137% | 1,349 | 1.35 | 1.30 |
| **7** | 976 | 170% | 1,658 | 181% | 1,769 | 1.06 | 0.98 |
| **8** | 1,014 | 199% | 2,020 | 232% | 2,350 | 0.90 | 0.77 |
| **9** | 1,011 | 176% | 1,783 | 230% | 2,324 | 1.02 | 0.77 |
| **10** | 1,045 | 193% | 2,021 | 287% | 2,998 | 0.93 | 0.62 |
| **Range** | | **89pp** | **930** | **171pp** | **1,824** | **0.54** | **0.90** |

*Table 1: Effect of sample boosting in lower-response deciles in the 2024 Time Use Survey*

Table 1 shows that actually we went slightly too far with the low-response decile boosting – we achieved most response in the lowest two response deciles (which would have had the lowest response without boosting) and relatively not quite enough response in the middle response deciles. However, the response with boosting is much more evenly distributed than it would have been without boosting, as reflected in the range of weights in the last two columns.

As an aside, this survey which only had online and telephone response options (no face-to-face interviewing) achieved a uniformly higher level of response than we anticipated, partly due to the success of the online diary design.

### *6.2 Sample enumeration*

The 2025 General Social Survey also was conducted using Computer Assisted Web Interviewing (CAWI) and Computer Assisted Telephone Interviewing (CATI) only.

It had a short enumeration period of only one month for the CAWI plus one month for CATI follow-up.

Midway through the first month of enumeration (May) we could see that the households that under most financial stress, as modelled with the PLIDA data, were disproportionally non-responders – by quite a lot compared to the remainder of the sample.

| GSS Response by Financial Stress Decile | | | | | | |
|---|---|---|---|---|---|---|
| | | Mid-May | | Final | | |
| Financial stress decile | Target | Target Rate | Relative weight | Target Rate | Relative weight | Improvement |
| ALL | 10,002 | 84.20% | 1.00 | 132.7% | 1.00 | - |
| 1 | 787 | 102.40% | 0.82 | 151.6% | 0.88 | 0.05 |
| 2 | 792 | 101.50% | 0.83 | 149.2% | 0.89 | 0.06 |
| 3 | 761 | 99.60% | 0.85 | 150.4% | 0.88 | 0.04 |
| 4 | 796 | 93.10% | 0.90 | 142.7% | 0.93 | 0.03 |
| 5 | 774 | 95.60% | 0.88 | 145.2% | 0.91 | 0.03 |
| 6 | 757 | 93.00% | 0.91 | 145.8% | 0.91 | 0.00 |
| 7 | 815 | 89.20% | 0.94 | 141.2% | 0.94 | 0.00 |
| 8 | 899 | 80.60% | 1.04 | 132.2% | 1.00 | 0.04 |
| 9 | 1,179 | 79.10% | 1.06 | 129.8% | 1.02 | 0.04 |
| 10 | 1,275 | 57.10% | 1.47 | 101.4% | 1.31 | 0.17 |
| 99 | 1,166 | 65.20% | 1.29 | 105.1% | 1.26 | 0.03 |

*Table 2: Response in the General Social Survey by Financial Stress Decile*[1]

We were able to target these households in our follow-up to reduce the extent of under-representation in the responding sample – see decile 10 in the table in the graph, which compares mid-enumeration status with end of enumeration status. The improvement from mid-enumeration was greatest for this decile and, while it was still the lowest response decile, it achieved the target (whereas the other deciles exceeded it).

### *6.3 Survey estimation*

At the time of writing, we did not have the results showing definitive, quantitative improvements to survey estimation from DIFFAR available. However, our experiments show that the gains of using this household-level information can be up to 20% compared to using auxiliary information at small area level.

Prior to DIFFAR being formalised, we were able to improve the quality of the 2021 Census Post-Enumeration Survey by incorporating PLIDA benchmarks of people receiving government benefits.

A second historical example is our Labour Force Survey (LFS). The LFS is a repeating survey (a selected household is in the sample for 8 consecutive months) so for many non-responders in a given month we have information about their previous response which can help. We monitored at this and made some adjustments to estimates during the COVID-19

---

[1] Decile '99' contains those households with unknown financial stress (no auxiliary data).

pandemic, when suddenly we couldn't use face to face interviewing to supplement the CATI and CAWI.

However, we know little or nothing about LFS respondents in their first month in sample (the incoming rotation group). If we have a particularly low response rate in any given month for any reason, across the whole sample (eg interviewer industrial action) or in some regions (a natural disaster), we can use the PLIDA information to make adjustments to overcome underrepresentation of any groups in the responding sample. For example, we know unemployed people tend to be disproportionately represented in the sample if the overall response rate drops, so we can use the PLIDA information on government unemployment benefit recipients as benchmarks to adjust for this.

**7. Conclusion**

Augmenting household survey frames with auxiliary data enable improvement can help reduce sample sizes and reduce the amount of bias and variation in sample estimates (particular those arising from increased or differential non-response).

The ABS' Data Improved Frames from the Address Register project demonstrates a feasible way of achieving this for countries that do not have access to a population register, but do have relevant administrative data.

**References**

Australian Bureau of Statistics (2020) ABS Address Register, Users' Guide | Australian Bureau of Statistics, ABS Website, accessed 31 August 2025.

Australian Bureau of Statistics (2023) Person Level Integrated Data Asset (PLIDA) | Australian Bureau of Statistics, ABS Website, accessed 31 August 2025.