Pre-analysis plan: family and domestic violence leave survey experiment

Policy problem

In 2023, national system employees in Australia received a new guaranteed entitlement: paid leave for people experiencing family and domestic violence (FDV). This was available for employees of large and medium businesses from 1 February 2023 and from 1 August 2023 for employees of small businesses. From September 2024, the entitlement will be extended to non-national system employees. All employees, including those who are casual and part-time, receive up to 10 full days of paid leave per year.

The legislation contains provisions for an independent statutory review to commence as soon as practicable from 1 February 2024. The legislation stipulates the review must consider the operation of the new leave entitlement, including the impact of the amendments on people experiencing family and domestic violence, small business and sole traders, and consider both quantitative and qualitative data.

BETA has been asked to create an evidence base, including both qualitative and quantitative data that will be provided to the independent reviewer. The focus of this evidence is to understand the impact that the legislated entitlement has had. The focus of this pre-analysis plan in our research into the impact that the legislation has had on the general working communities attitudes towards the paid FDV leave and those who take it.

Examining the general working community attitudes is important. There is a complex relationship between legislation, attitudes and behaviour. Preventing FDV involves changing attitudes, and evidence indicates that introducing legislation can alter attitudes. If the leave entitlement has reduced social stigma and strengthened support for those who have experience FDV, then victims may be more likely to speak up and access the leave or other support, increasing the intended effectiveness of the leave.

Trial Aims

To investigate whether awareness of the new Family and Domestic Violence (FDV) paid leave legislation and provisions reduces stigmatising attitudes towards those who have experienced FDV and take the paid FDV leave.

Trial design

This is a two arm Randomised Control Trial (RCT) involving an online survey experiment with members of the National Employment Standards (NES) working population. The experiment aims to leverage low awareness of the new Australian paid FDV leave legislation among the working population. Half our participants will learn about the legislation (through educational feedback) *before* responding to a vignette that requires participants to make judgements about a fictional person experiencing FDV and taking leave from work. Participants will also answer general attitudinal questions about FDV (Treatment condition). The other half of our participants will learn about the legislative requirements *after* they have answered questions associated with the vignette and their attitudes towards FDV (Control condition).

We will recruit a sample of participants who are representative of the Australian general working population through Pure Profile's research participant panel. The recruitment flow, eligibility, and exclusion criteria for the RCT survey experiment is outlined in Figure 1.

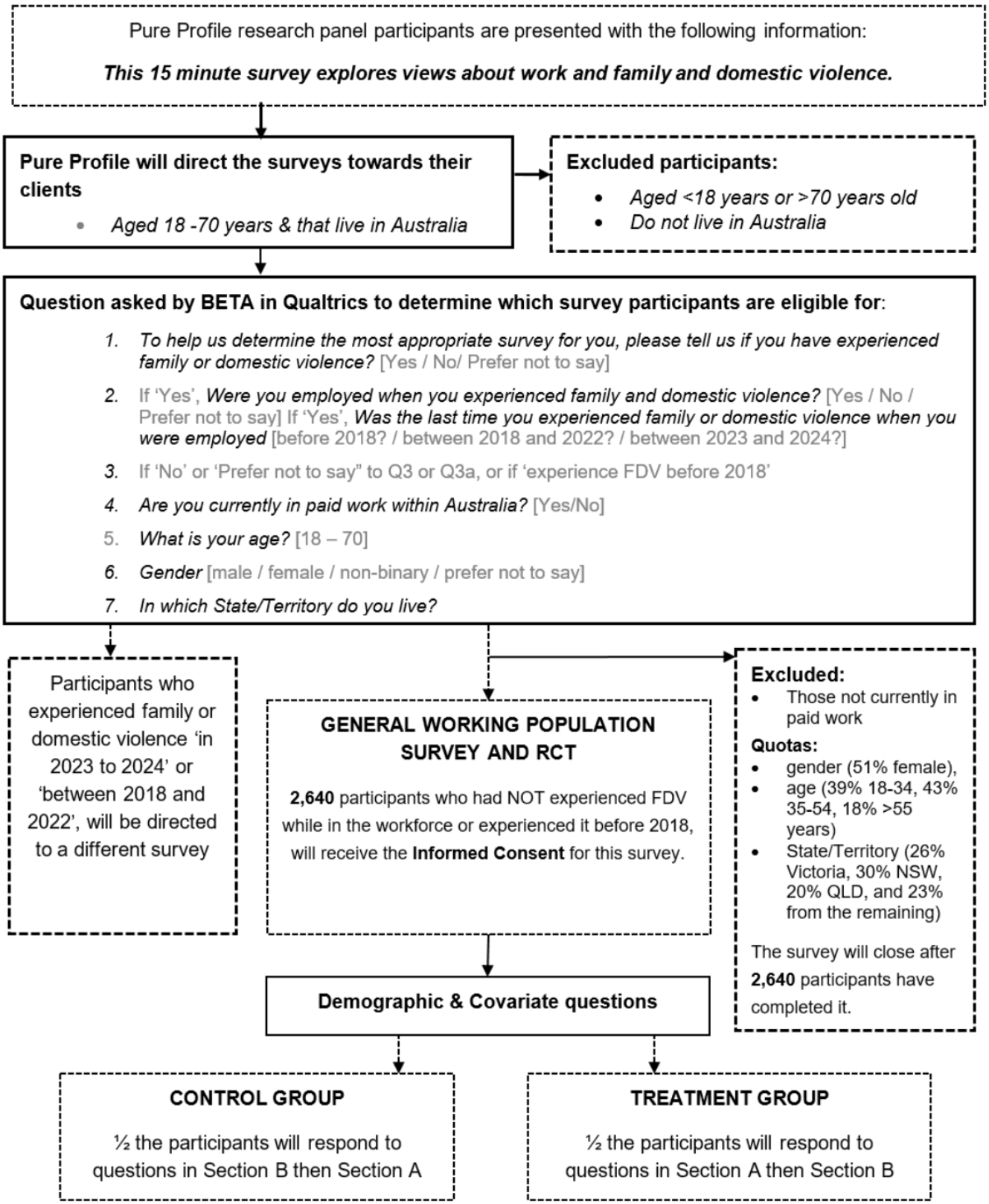
We will use hard quotas with the following targets:

- Gender – 51 per cent female
- Age – 39 per cent aged 18 to 34, 43 per cent aged 35 to 54 and 18 per cent over 55.
- State or territory – 26 per cent from Victoria, 30 per cent from New South Wales, 20 per cent from Queensland and 23 per cent from the remaining states and territories

We will use soft quotas for the following targets:

- CALD status – approximately 22 per cent who either mostly speak a language other than English at home and/or were born overseas.
- Employment status – Approximately 20 per cent who are employed on a casual basis.

Figure 1. Participant recruitment flow



Interventions

We will present an activity to participants to encourage learning about specific aspects of the legislation. It will include a quiz on features of the legislation including administrative requirements, and privacy protections. Participants will receive feedback on correct answers in real time to increase their knowledge of the legislation. We will randomise participants into two arms – those who receive educational feedback about the legislative requirements *before* we measure attitudes and stigma towards those who experience FDV and take the paid FDV leave (treatment) and those who receive the educational feedback *after* we measure their attitudes and stigma (control).

Outcome measures

There are two sets of outcome measures. The first set will be measured in response to a vignette. The vignette describes an employee who experiences domestic violence and takes work leave to deal with this. Outcomes measures relate directly to perception of the employee. The second set are general attitudinal outcomes towards those who experience FDV.

We explored the use of one of the ANROWS outcome measures as a primary outcome. However, attitudes are a more distal measure of our intervention compared to the discrimination outcome measures. Hence we relegated it to secondary measures. Moreover, there is no existing data available for power calculations and we were concerned about the scale's sensitivity to change. We cannot also tell what the score changes mean in terms of actual attitude changes. We made the competence discrimination scale as a secondary outcome for the same reason.

Primary outcomes

Discrimination outcome measures

- Bonus (numeric). In the vignette participants are informed that employees may receive a bonus between $0 and $1000. Participants will be asked what bonus is appropriate for the employee in whole dollars between 0 and 1000.
- Management (binary). Participants are asked whether the employee should be assessed for a management position in the next year (1 = yes, 0 = not yes)

We will undertake missing value imputation (described later) if more than 10% of the responses are missing OR if there is differential attrition between treatment and control groups.

Secondary outcomes

For each of these outcomes there is a small chance that there will be missing responses as the items will not be compulsory. For each scale, we will compute a score for participants who respond to half of the items in the scale. Otherwise their total score will be marked as missing. Thresholds are described below.

Discrimination outcome measure

- Competence scale (numeric). This will be measured using 6 items, each with a 10-point sliding response scale. The outcome will be the mean of these responses. The mean will be calculated where participants have responded to at least 3 of the 6 items. Participants who respond to only 2 or fewer items will not have a score calculated and will be marked as missing data.

Attitude Outcome measures

- ANROWS Attitudes Towards Gender Inequality Scale (adapted, AGIS scale). Five items measuring attitudes towards domestic violence. Participants respond on a 4-point sliding scale (Disagree = 0, Slightly agree = 1, Agree = 2, Strongly agree =3). The outcome will be constructed by calculating the mean of responses. We will calculate the outcome if participants complete at least 4 responses. For participants who respond to fewer than 3 responses, this outcome will be marked as missing.
- Twelve items measuring attitudes towards domestic violence (DVS scale). Participants respond on a 4-point sliding scale (Disagree = 0, Slightly agree = 1, Agree = 2, Strongly agree =3). The outcome will be constructed by calculating the mean of responses if participants complete at least 6 responses. For participants who respond to fewer than 6 responses, this outcome will be marked as missing.

We will not undertake missing value imputation/modelling for secondary outcome measures.

Hypotheses

Primary

H1: Mean bonus amount will be higher in the treatment arm as compared with the control arm (treatment $>$ control).

H2: There will be a higher proportion of people recommending assessment for management for the fictional employee in the treatment arm as compared with the control arm (treatment $>$ control).

Secondary

H1: Mean competence scores will be higher in the treatment arm as compared with the control arm (treatment $>$ control).

H2: Mean DVS scores will be lower in the treatment arm as compared with the control arm (treatment $<$ control).

H3: Mean AGIS scores will be lower in the treatment arm as compared with the control arm (treatment $<$ control).

Randomisation

Participants will be individually randomised using the Qualtrics platform. We will randomise with equal probability to each arm.

Sample size and power calculations

With 1,320 participants per arm, we will have 90% power to detect an effect of:

- $25 higher in the mean recommended bonus in treatment compared to the control group
- 5 percentage point higher in proportion recommending fictitious person for management (i.e. from 50% to 55%) in the treatment group compared to the control group

We will use an alpha level of 10% and 90% power. We have chosen these settings because the intervention is extremely low risk and it would be worse to reject a possible real effect than to accept a possibly spurious one.

Method of analysis

The principal analysis of the effect of the intervention will consist of a covariate-adjusted comparison of our primary outcomes. This estimate, confidence intervals and p-values will be derived from a linear regression model using robust (HC2) standard errors and with the following specification:

$$Y_i = \beta_0 + \beta_1 Z_i + \beta_2 X_i + \beta_3 Z_i X_i + \in_i$$

Where i is an index for each individual in the trial, Y is the individuals score on the outcome measure, β_0 is the intercept, Z is a treatment assignment indicator, β_1 is a coefficient representing the average treatment effect for the intervention relative to control, X is a vector of four mean centred covariates (see Covariates section below), and ZX is the interaction of the treatment indicator vector with the mean centred covariate indicator vector and $\in$ is the individual error term.

All primary outcomes have directional hypotheses so we will do a one-sided test. Secondary outcomes will also involve one-sided tests.

We will also perform exploratory subgroup analyses by factors such as public or private employment sector and highly gendered industries.

Covariates

We will adjust for the following pre-randomisation variables (collected prior to randomisation) in our regression:

- Male (binary variable, 1 = male, 0 = not male)
- Education (binary variable, 1 = has tertiary education, 0 = does not have tertiary education)
- CALD (binary variable, 1 = born outside Australia, and home language was a language other than English, 0 = born in Australia and/or home language was English)
- Prior awareness of the new paid leave entitlement (binary variable, 1 = did NOT identify that there was a leave that casuals could access (NOT aware), 0 = identified that there was a leave that casuals could access (AWARE)

Trial threats

Missing data is a plausible threat to this trial. If participants fail to respond or select 'I don't know' to a large proportion of items there will be a reduction in power. It is also possible that missingness will be related to the arm to which participants have been assigned. For example if the intervention reduces stigma and increases compassion these participants may be more willing to engage with the outcomes. To address this we will take a multi-stage approach.

1 We will check missingness in primary outcomes while data is being collected as part of our data quality assurance assessment. If missingness is more than 10%, we will recruit additional sample where possible.

2 At the end of data collection, we will examine the rate of missingness in our primary outcome variables. If this is below 10% in both arms we will conduct a complete case analysis.

3 If the rate of missingness is above 10% in at least one arm, we will examine the pattern of missingness by first checking if there is differential attrition and examine the pattern of missinginess. If there is evidence that the missingness is completely at random (MCAR) we will conduct a complete case analysis. We may boost the sample to ensure power is maintained.

4 If the missingness is missing at random (MAR) or missing not at random (MNAR[1]), we will conduct Lee bounds[2] and present both complete case analysis and Lee bounds.

The trial also artificially increases knowledge of the relevant legislation and its provisions. In a field setting it would be extremely difficult to achieve this level of engagement. However, the aim is to assess whether over time, as knowledge of the provisions in the legislation become more widely known, is there likely to be a reduction in stigmatising attitudes? This trial will not provide direct evidence for the magnitude of such a reduction due to its artificial setting.

Interpretation and reporting

For our primary hypotheses, we will use null hypothesis statistical testing in order to facilitate decision making about whether to treat an effect as real. However, we will also make use of non-significant primary analyses, secondary and exploratory analyses in order to provide context and to highlight interesting avenues for further research. We will clearly delineate these analyses when communicating findings.

We will accept that the intervention is effective, if any of the two primary outcomes are statistically significant. Therefore we will adjust alpha using the Holm[3] method.

---

[1]There is no way to definitely test for MNAR but it is a possibility that missing data will be related to treatment assignment in this experiment so we will proceed with analyses for MNAR if we find no evidence for MCAR or MAR.

2 Lee, DS (2002) 'Trimming for bounds on treatment effects with missing outcomes', *Centre for Labor Economics, Working Paper 51*.

3 Rubin, M (2021) 'When to adjust alpha during multiple testing: A consideration of disjunction, conjunction, and individual testing', *Synthese,* 199:10969-11000, doi: 10.1007/s11229-021-03276-4

When reporting proportions treatment effects, standard errors, and confidence intervals will be presented as percentage point differences. Absolute p-values will be reported. We will provide these outputs for all primary hypotheses and any pre-specified secondary analyses.

Pre-analysis plan commitments

We have two standard commitments:

- No analysis has been undertaken prior to the completion of this pre-analysis plan. This exclude checking for data quality and missingness.
- We will be transparent about, and provide justification for, any deviations (additions or omissions) from this plan.