



Australian Government



Services
Australia

Evaluation Report

Pre-Call SMS Trial

September 2025

© Commonwealth of Australia 2026

ISBN: 978-1-923278-56-1

This publication is available for your use under a [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/legalcode) licence, with the exception of the Commonwealth Coat of Arms, the Services Australia logo, photographs, images, third party materials, materials protected by a trademark, signatures and where otherwise stated. The full licence terms are available from creativecommons.org/licenses/by/4.0/legalcode.



Use of report material under a [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/legalcode) licence requires you to attribute the work (but not in any way that suggests that the Services Australia or Treasury endorses you or your use of the work).

Material used 'as supplied'

Provided you have not modified or transformed report material in any way including, for example, by changing the text; calculating percentage changes; graphing or charting data; or deriving new statistics from published statistics – then Services Australia and Treasury prefer the following attribution:

Source: *The Commonwealth of Australia 2025 'Evaluation Report: Pre-Call SMS Trial'*

Derivative material

If you have modified or transformed report material, or derived new material from those of the report in any way, then Services Australia and Treasury prefer the following attribution:

Based on Commonwealth of Australia data.

Use of the Coat of Arms

The terms under which the Coat of Arms can be used are set out on the Department of the Prime Minister and Cabinet website (see www.pmc.gov.au/government/commonwealth-coat-arms).

Other uses

Enquiries regarding this licence and any other use of this document are welcome at:

Manager
Media Unit
The Treasury
Langton Crescent
Parkes ACT 2600
Email: media@treasury.gov.au

In the spirit of reconciliation, the Services Australia and Treasury acknowledge the Traditional Custodians of country throughout Australia and their connections to land, sea and community. We pay our respect to their Elders past and present and extend that respect to all Aboriginal and Torres Strait Islander peoples.

Services Australia led this trial with advice from the Australian Centre for Evaluation (ACE), in the Australian Treasury, on the evaluation design and reporting. For this reason, it is being published on the ACE web site.

Contents

Executive Summary	4
Evaluation Approach	4
Key Findings	4
Conclusions and recommendations	4
Background	5
Methodology	5
Implementation	5
Results	6
Impact of SMS on Outbound Contact Success	7
Impact of Revising SMS Language on Outbound Contact Success	8
Staff Survey Results	9
Conclusions.....	10
Limitations.....	10
Recommendations	10
Appendix A: Additional Calculations and Results.....	11
Appendix B: Key Concepts.....	18
Appendix C: Previous Research	19
Appendix D: Technical Details	20
References	25

Executive Summary

Services Australia (the agency) delivers government payments and services. To support this work, the agency frequently contacts customers by phone to obtain additional information for a number of reasons, such as to finalise a claim. A 'Pre-Call SMS' is often sent to alert customers of the upcoming call. However, there has been limited evidence about the effectiveness of the Pre-Call SMS in prompting customers to answer the phone.

Evaluation Approach

A Randomised Controlled Trial (RCT) was used to test the effectiveness of the Pre-Call SMS and explore whether adjusting the wording improves customer call answer rates. The trial was delivered over a five-week period by staff across two offices, primarily involving teams focused on Families (for example Family Tax Benefit, Child Care Subsidy) and Students (for example Youth Allowance, Austudy) related work.

Key Findings

Sending a Pre-Call SMS increased customer call answer rates. Customers who received a Pre-Call SMS were 12 percentage points more likely to answer the call compared to those who did not receive a Pre-Call SMS. The proportion of successful outbound calls with a Pre-Call SMS was 64%, compared to 52% with no SMS. This was a statistically significant difference.

Updating the SMS wording did not have a meaningful difference on customer call answer rates. The proportion of successful outbound calls with the existing Pre-Call SMS was 66% compared to 63% for the revised SMS; this difference was not statistically significant.

Conclusions and recommendations

The results of this RCT indicate that sending a Pre-Call SMS improves outbound contact success, and that minor adjustments to wording made no meaningful difference. These outcomes align with previous observational and trial-based research on Pre-Call SMS effectiveness. Successful completion of this RCT establishes the feasibility of undertaking further trials across the agency in future:

1. Continue sending a Pre-Call SMS with the existing wording to alert customers to an upcoming phone call. This will provide better staff and customer experience with currently available technology
2. Consider the net benefit of rolling out Pre-Call SMS in other business areas, where contact rates may increase by up to 12 percentage points, noting that program level variations can occur
3. Explore alternate technologies and approaches to further enhance outbound contact experiences for customers and staff
4. Use the experience, knowledge and resources developed during this pilot to run further evaluations and trials that support evidence-based decision making across the agency.

Background

To deliver government payments and services, the agency frequently contacts customers by phone to obtain additional information for a number of reasons, such as to finalise a claim. A 'Pre-Call SMS' is often sent to alert customers of the upcoming call.

The purpose of this evaluation was to test whether sending a Pre-Call SMS increases customer call answer rates, and whether changes to the SMS wording further improve call answer rates.

Previous analysis of the agency Pre-Call SMS data used a quasi-experimental design to associate a Pre-Call SMS with a 5.5 percentage point increase in customer call answer rates. This evaluation aims to provide stronger evidence about the Pre-Call SMS through the use of a randomised controlled trial (RCT) by providing a more reliable counterfactual.

This RCT supports decision-making regarding processes and guidance for outbound calls. In addition, the RCT was also deliberately designed to build capability in delivering impact evaluations across the Agency, and to embed impact evaluation across the agency moving forward.

Methodology

This evaluation used a RCT to test the effectiveness of sending a Pre-Call SMS on outbound call success (customer call answer rates). Outbound call success was defined as a call that staff make to a customer, where:

- The customer answered and
- The call lasted for at least 30 seconds.

The RCT included Centrelink customers who required an outbound call from the Families and Students program areas, as well as other general work allocated to the staff participating.

Implementation

The RCT was carried out in two Smart Centres from 16/07/2025 to 24/08/2025 inclusive.

Service Officers at the participating offices were randomly assigned to one of the following groups:

Table 1: Trial Arm descriptions

Trial Arm	SMS Sender ID	SMS Wording
Send the existing SMS	Centrelink	A Services Australia Service Officer will call you today from a private number. Please answer this call. Do not reply to SMS
Send a revised SMS	Centrelink	We will call you today from a private number with some important information. Please answer this call. Do not reply by SMS
Do not send an SMS prior to making an outbound call	N/A – no SMS sent	N/A – no SMS sent

The RCT was reviewed by the agency's internal data and ethics risk assessment processes.

At the completion of the trial, a voluntary survey was emailed to staff who had participated. The purpose of the survey was to explore staff perspectives, sentiment and feedback on both the Pre-Call SMS and their participation in the trial.

Results

The trial provided 4,833 total observations for analysis.

Table 2: Summary of Trial Observations

Trial Arm	Observations (N)	Teams (Clusters)	Service Officers
Existing SMS	1,817	7	93
Revised SMS	1,528	9	104
No SMS Sent	1,488	7	78
Total	4,833	23	275

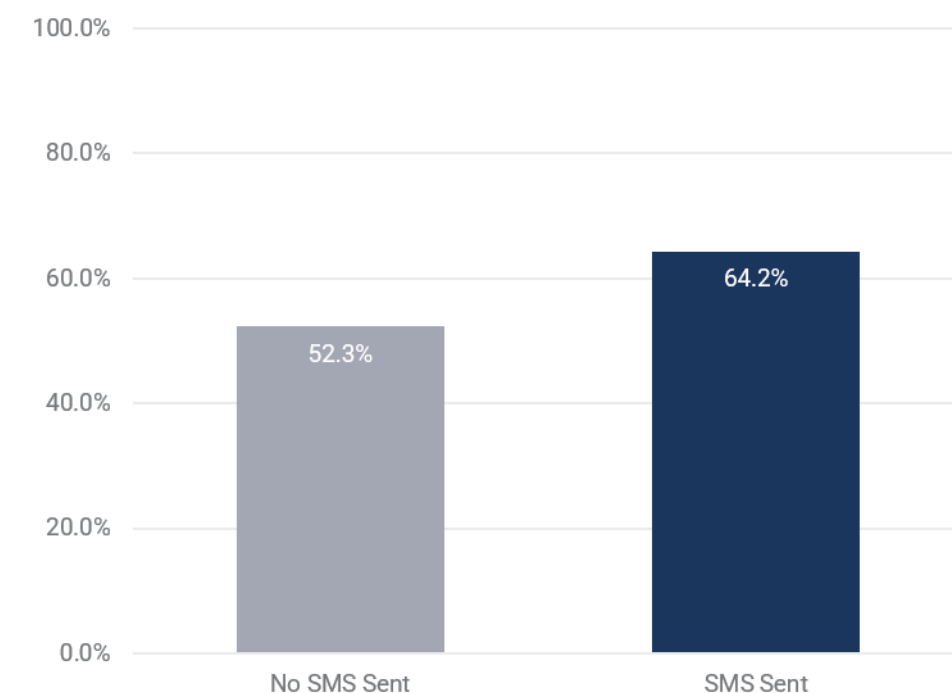
Impact of SMS on Outbound Contact Success

Pre-Call SMS Effectiveness

Sending any Pre-Call SMS to a customer prior to an outbound contact increased the overall call success rate by 12 percentage points, from 52% without a Pre-Call SMS to 64% with a Pre-Call SMS¹. This was a statistically significant increase in successful outbound contacts due to the Pre-Call SMS.

Sending the existing Pre-Call SMS increased the success rate to 65%, while the adjusted version increased it to 63%, compared with 52% without an SMS. Both variations delivered statistically significant increases in successful outbound contacts.

Figure 1: Impact of Pre-Call SMS on Outbound Contact Success



Source: Pre-Call SMS and outbound contact metadata.

This figure shows the baseline (52%) and SMS effects (64%) from a linear regression model of outbound contact success, by whether or not the participating service officers were directed to send a Pre-Call SMS². The difference between treatment and baseline was statistically significant ($p < 0.001$). $N = 4,833$. Clusters = 23.

Pre-Call SMS Effectiveness for Families and Students cohorts

A Families-related Pre-Call SMS increased the success rate to 66%, compared to 53% without an SMS. A Students-related Pre-Call SMS increased the success rate to 60%, compared to 50% without an SMS. The Families-related result is statistically significant. Both Families and Students improvements are meaningful, though the limited number of Students teams means a statistical test was not performed.

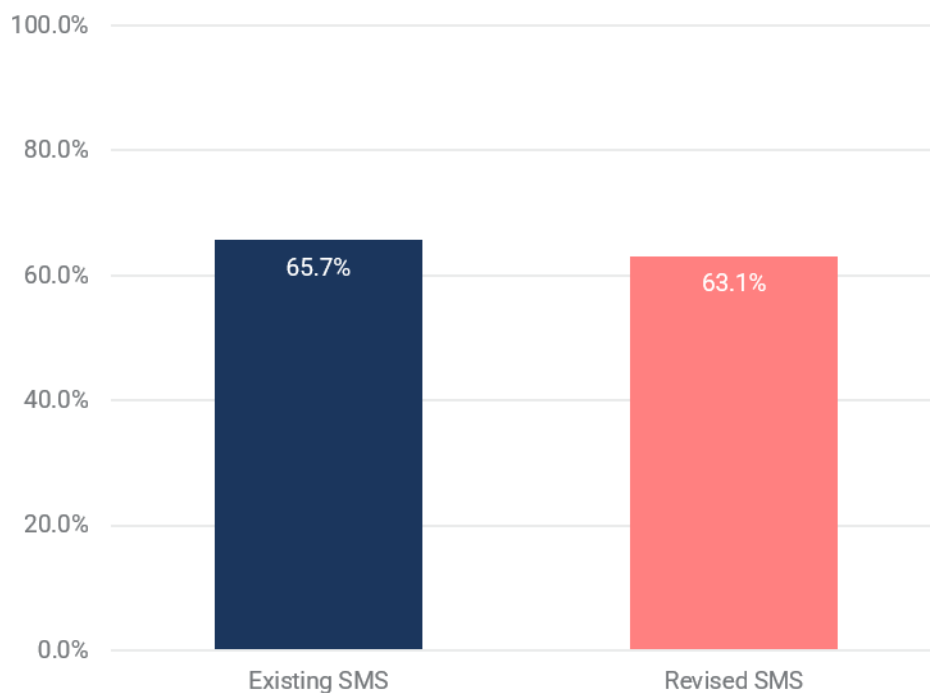
¹ Results were consistent between linear and logistic regression model specifications.

² University of Illinois (n.d.). Statistical methods are discussed in *Appendix D: Technical Details*.

Impact of Revising SMS Language on Outbound Contact Success

There was no meaningful difference between the Pre-Call SMS types. There was a minor decrease in success for the revised Pre-Call SMS, compared to the existing Pre-Call SMS. However, this could be due to natural variation³.

Figure 2: Impact of Revised SMS Language on Successful Outbound Contact



Source: Pre-Call SMS and outbound contact metadata.

This figure shows the SMS effects from a linear regression model of outbound contact success, by each Pre-Call SMS certain service officers were directed to send. N=3,345. Clusters=16.

³ Further discussion is provided in *Appendix D*.

Staff Survey Results

Responses were collected separately from service officers and team leaders.

Table 3: Number of responses by staff group

Staff Group	Responses	Response rate
Service Officer	60	22%
Team Leader	9	39%

Overall service officer sentiment towards the pilot was positive, with 62% of service officers saying it was a positive experience, and 28% neutral. Team leaders had a higher positive sentiment at 78%. 71% of service officers expressed they would like to participate in future pilots, and 77% of service officers thought the instructions provided to them were helpful.

Service officer comments were consistent with the data from the trial results, with service officers not sending the SMS reporting they had increased difficulty progressing their calls with customers, as seen in the following example comments:

- “Overall, I had less successful contact attempts with customers but if call processes are followed correctly (voicemail left, request for information sent) correct outcomes were still achieved for the customer with only a slight delay. Whether that was the (customer) calling back or responding to the request for information.” (No SMS Trial arm)
- “I personally found using the alternative/ approved SMS messaging did not have any real (effect) on the likelihood of the customer answering the call. That is to say, I don’t feel it made much difference.” (new SMS Trial arm)
- “A lot of (customers) were not happy that the Pre-Call SMS were not sent. They expressed feelings of concern that the call could be a scam. This indicates that customers do value the use of Pre-Call SMS, and it does reassure the customers that we are who we say we are.” (No SMS Trial arm)

While some staff reported that customers asked about not receiving a Pre-Call SMS, a review of complaints received during the trial period found no formal customer complaints on this issue.

Conclusions

This RCT indicates the Pre-Call SMS is effective at increasing customer call answer rates. One possible explanation is that the SMS helped customers anticipate the call and be prepared to answer.

The results reflected a 12 percentage point difference between the no SMS and SMS groups, which is larger than the previous quasi-experimental design analysis which reflected a 5.5 percentage point increase. This may reflect the trial environment and difference in SMS responses across different Centrelink work types. For example, the Families cohort in this trial showed a higher successful customer call answer rate (approximately 10%).

The consistency with prior research validates the continued use of the Pre-Call SMS, with observational data to continue to track effectiveness. The successful completion of the RCT also establishes the feasibility of further trials on other agency processes and functions.

There was no meaningful difference between the existing and revised SMS wording. This is likely due to the limited changes made to the adjusted SMS wording.

Limitations

The RCT was performed in Centrelink service delivery in two offices whereby teams were predominately performing Families and Students-related work due to RCT feasibility and logistics. Due to this, results may differ across different program areas or sites.

Privacy, technology and correspondence standards limited the extent to which the SMS wording could be altered. This may have contributed to the lack of difference found between the two messages. Future evaluations could consider testing more substantial changes to the Pre-Call SMS wording or other potential process changes.

Staff were aware they were participating in a trial which may have led to closer adherence to instructions regarding call behaviours. If instructions are followed, the measure of success is driven by customer response rather than staff actions, so the impact of staff awareness of the trial should have been minor.

Recommendations

- a. Continue sending a Pre-Call SMS with the existing wording to alert customers to an upcoming phone call. This will provide better staff and customer experience with available technology.
- b. Consider the net benefit of rolling out Pre-Call SMS in other business areas, where contact rates may increase by up to 12 percentage points, noting that program level variations can occur. Although Pre-Call SMS is broadly used, there may be further opportunities to use it for other business processes when appropriate.
- c. Explore alternate technologies and approaches to further enhance outbound contact experiences for customers and staff. The remaining percentage of unsuccessful outbound contacts (36%) represent missed opportunities for customers and staff to engage.
- d. Use the experience, knowledge and resources developed during this pilot to run further evaluations and trials that support evidence-based decision making across the agency. Overall, the staff response to the trial was positive and the trial was delivered with minimal impact to service delivery operations.

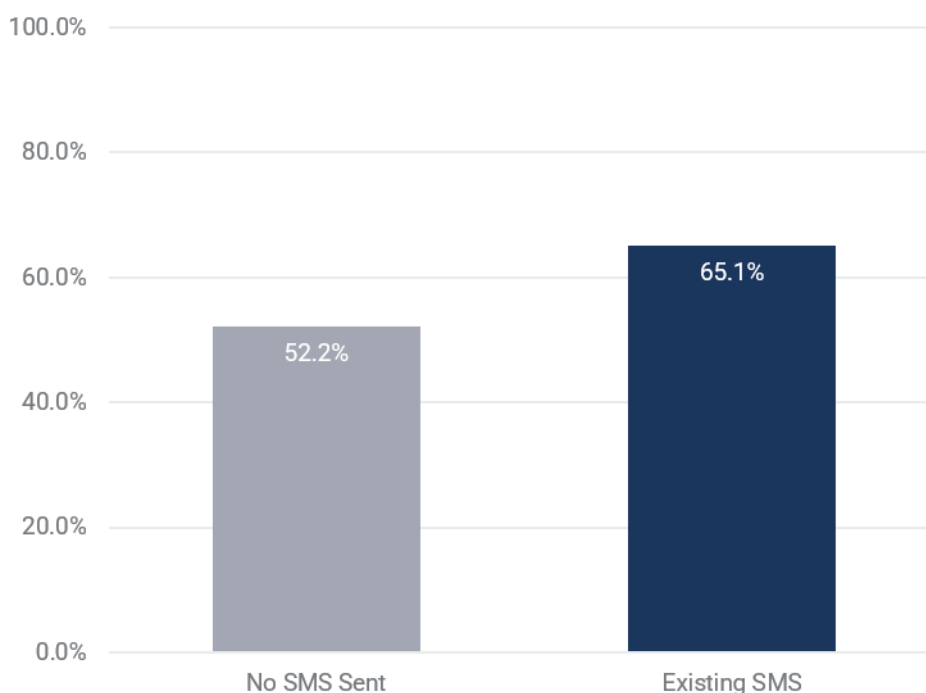
Appendix A: Additional Calculations and Results

Additional Results

Comparing Existing SMS to Baseline

Given the pooled primary outcome is statistically significant, a derivative outcome is whether the existing SMS is different to not sending an SMS.

Figure 3: Impact of Existing Pre-Call SMS on Outbound Contact Success

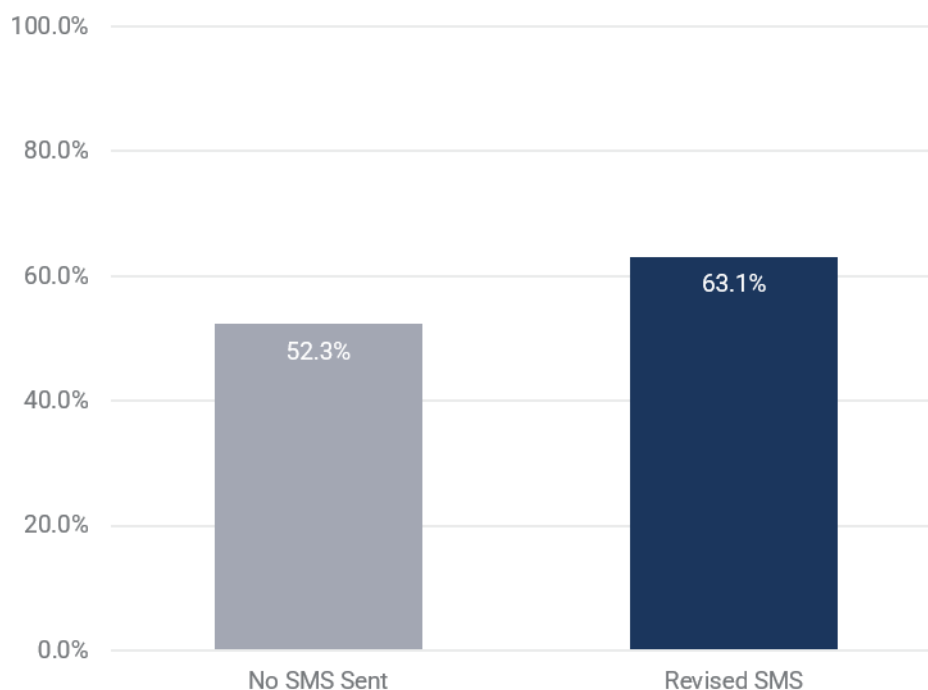


This figure shows the baseline and existing SMS effect from a linear regression model of outbound contact success, by whether or not the participating service officers were directed to send the existing Pre-Call SMS. The difference between treatment and baseline was statistically significant ($p < 0.001$). $N = 3,305$. Clusters = 14.

Comparing Revised SMS to Baseline

Given the primary outcome is statistically significant, a derivative outcome is whether the revised SMS is different to not sending an SMS.

Figure 4: Impact of Revised Pre-Call SMS on Outbound Contact Success

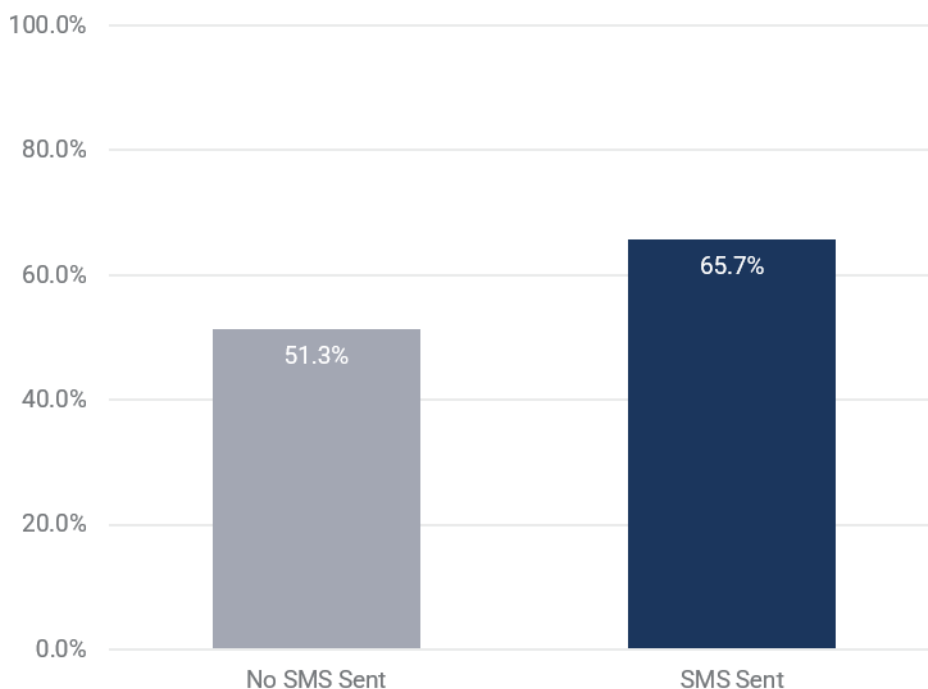


This figure shows the baseline and revised SMS effect from a linear regression model of outbound contact success, by whether or not the participating service officers were directed to send the revised Pre-Call SMS. The difference between treatment and baseline was statistically significant ($p < 0.001$). $N = 3,016$. Clusters = 16.

Per-Protocol (PP)

A secondary outcome of interest is the subset of primary analysis records where service officers acted according to their trial arm 'per-protocol' (PP). This is a secondary outcome as it no longer assumes randomisation⁴. This is unlike the primary outcome, which focuses on the 'intent-to-treat (ITT)', therefore assuming randomisation.

Figure 5: Impact of Per-Protocol Pre-Call SMS on Outbound Contact Success



This figure shows the baseline and SMS effect from a linear regression model of outbound contact success, by whether or not the participating service officers were directed to send an SMS and acted according to this direction. N=3,930. Clusters=23.

Despite the loss of 903 observations, the results remain stable between ITT and PP model specifications. The marginal average baseline decreased by 1 percentage point and the marginal average treatment effect increased by 2 percentage points, consistent with expectations of ITT providing a more conservative (and accurate) result⁵.

The primary outcome of interest applied an 'intent-to-treat' (ITT) trial design, where results are generated in line with randomisation, thus preserving the benefits of randomisation.

Data limitations are expected with the required complex data curation. Service officers assigned to the adjusted or no SMS groups while sending the existing Pre-Call SMS may indicate out-of-scope work that was included in model data. Some service officers may not have acted according to their trial arm due to misunderstandings about what was required as part of the trial.

⁴ McCoy, C. E. (2017).

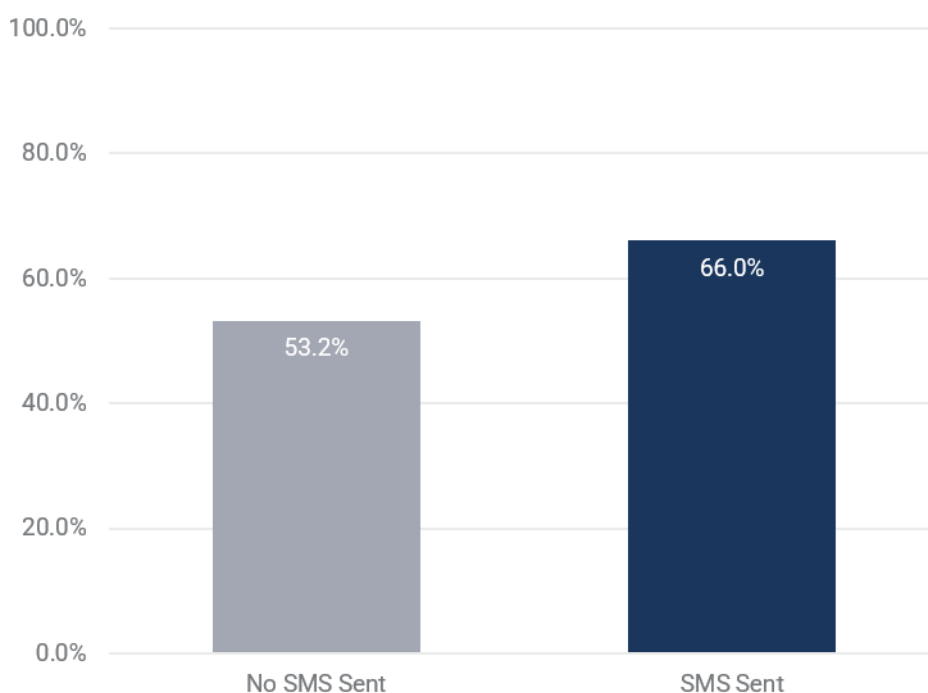
⁵ ibid.

Cohort Analysis

A secondary outcome of interest is how Families or Students subgroups influenced outbound contact success.

Families

Figure 6A: Impact of Families-Related Pre-Call SMS on Outbound Contact Success

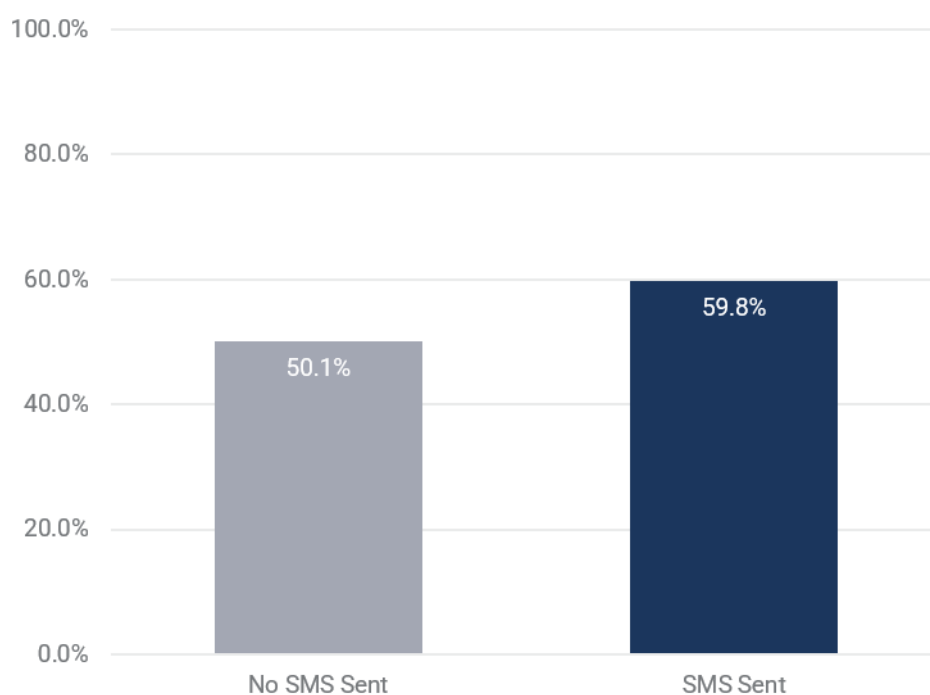


This figure shows the baseline and SMS effect from a linear regression model of outbound contact success, by whether or not the participating Families service officers were directed to send a Pre-Call SMS. N=3,470. Clusters=20.

When limiting to Families data, the marginal average baseline increased by one percentage point and the SMS effect increased by two percentage points. Families service officers contributed 20/23 teams and 72% of outcomes, so results resemble those from the primary outcome.

Students

Figure 6B: Impact of Students-Related Pre-Call SMS on Outbound Contact Success



This figure shows the baseline and SMS effect from a linear regression model of outbound contact success, by whether or not the participating Students service officers were directed to send a Pre-Call SMS. N=1,363. Clusters=3.

When limiting to Students data, the marginal average baseline decreased by 2 percentage points and the SMS effect decreased by 4 percentage points.

Accessible Graph Data

Table 4: Impact of SMS

Trial Arm	N	Clusters	Success (%)	Difference from Control	Standard Error
No SMS Sent	1,488	7	52.3028	NA	NA
SMS Sent	3,345	16	64.2486	11.9458	1.298

Table 5: Impact of Existing SMS

Trial Arm	N	Clusters	Success (%)	Difference from Control	Standard Error
No SMS Sent	1,488	7	52.1561	NA	NA
Existing SMS	1,817	7	65.1313	12.9753	0.7768

Table 6: Impact of Revised SMS

Trial Arm	N	Clusters	Success (%)	Difference from Control	Standard Error
No SMS Sent	1,488	7	52.2559	NA	NA
Revised SMS	1,528	9	63.0591	10.8032	1.982

Table 7: Impact of Revised SMS Language

Trial Arm	N	Clusters	Success (%)	Difference from Control	Standard Error
Existing SMS	1,817	7	65.6508	NA	NA
Revised SMS	1,528	9	63.1181	-2.5327	2.25

Note: success calculations are ‘marginal’ or ‘unconditional’ estimates, which can vary based on what subset of data was used (see *Statistical Inference in Appendix D: Technical Details*). Existing SMS estimates are affected by rounding error at 0 decimal places.

Table 8: Per-Protocol Differences by Trial Arm

Trial Arm	N	Clusters	Success (%)	Difference from Control	Standard Error
No SMS Sent	1,331	7	51.3372	N/A	N/A
SMS Sent	2,599	16	65.7385	14.4013	N/A

Note: refer to the Additional Results > Per-Protocol section. Standard errors were not calculated; Per-Protocol does not assume randomisation and associated standard errors are not used in the report.

Table 9A: Impact of Families-Related SMS

Trial Arm	N	Clusters	Success (%)	Difference from Control	Standard Error
No SMS Sent	953	6	53.1706	N/A	N/A
SMS Sent	2,517	14	66.0028	12.8321	1.795

Table 9B: Impact of Students-Related SMS

Trial Arm	N	Clusters	Success (%)	Difference from Control	Standard Error
No SMS Sent	535	1	50.0935	N/A	N/A
SMS Sent	828	2	59.7826	9.6892	0.787

Appendix B: Key Concepts

The agency sends Pre-Call SMS and stores metadata of those actions in datasets. Similarly, the agency makes outbound contacts and stores metadata in datasets. There is no intrinsic relationship between Pre-Call SMS and outbound contacts; each are sent in different parts of the overall processing environment, and one does not necessitate the other. To perform this RCT, assertions about their relationship were required.

Success Definition

The *Methods* section gave the following definition of success: “a successful outbound contact occurred when a service officer and customer persisted with the contact for at least 30 seconds.”

This 30 second threshold, while somewhat arbitrary, removes calls which connected to a voicemail message rather than a customer and very short calls which could not have been successful.

Multiple Contacts and Multiple Pre-Call SMS

Multiple contacts and multiple SMS argued to be related into a single overall attempt were aggregated. This resulted in a single row in the model dataset for what may have been multiple Pre-Call SMS sent, and multiple outbound contacts made. This aggregation aligns data to the process, prevents double-counting, ensures comparability across process outcomes for the same work and minimises analyst judgment in constructing metrics.

For example, one application of the process may lead to:

- The first individual attempt leading to a successful contact
- The third individual attempt using the customer’s first phone number for the second time leading to a successful contact, after unsuccessful contact using each of the phone numbers
- The second individual attempt on the second day using the customer’s one phone number, leading to an unsuccessful contact

This method addresses the problem with associating individual outbound contacts and individual Pre-Call SMS. An ‘outbound attempt’ can be characterised by multiple individual outbound attempts and multiple Pre-Call SMS. The process does not explicitly restrict service officers to one Pre-Call SMS per customer per day, though this restriction would be consistent with the standard Pre-Call SMS referring to a contact ‘today’.

For example, service officers may send a Pre-Call SMS:

- for every individual outbound attempt
- on back-to-back days

It’s possible that customers are serviced by different service officers on the same day, also possibly leading to multiple Pre-Call SMS. The analyst needs to navigate the ambiguity of process and practice, either by generalising to overall ‘outbound attempt’ or making erroneous judgements about a wide range of potential relationships between Pre-Call SMS and outbound contacts. The former was chosen.

Appendix C: Previous Research

Previous research informed the trial in the following ways:

- **Contestability:** SMS improves outbound contact success, both in the initial pick-up and in the desired business outcomes, measured to be between 5.5 and 13.5 percentage points

Observational Analysis

A previous report in 2024 **identified a 5.5 percentage point improvement, from 56.7% to 62.2%**. The analysis provided was observational, summarising the relationship between Pre-Call SMS and outbound contact success, but not necessarily accounting for all the complexities of service delivery operations. This motivated the call for a randomised control trial, which better accounts for those complexities. The coverage included a basic breakdown of outbound contact success, based on whether or not a Pre-Call SMS was sent, and further broken down by program area, Smart Centre and staff area, and what customers say when they complain about the Pre-Call SMS.

Appendix D: Technical Details

Trial Design

The RCT followed a *parallel design*, where service officers were assigned to a single trial arm for its duration.

There were multiple conceptual frameworks. For testing the effect of sending either SMS, it was *superiority*, whether the SMS *improved* success compared to not sending the SMS. For testing the effect between SMS, it was *equivalency*, whether there was a *difference* between sending either SMS.

The unit of randomisation were clusters of service officers at the team level. Teams have unique identifiers and were constituted by APS4 service officers.

The allocation ratio for each of the three trial arms was 1:1:1. The two SMS trial arms were pooled for testing the effect of sending either SMS, giving an allocation ratio of 1:2. There were minor differences in the number of teams in each trial arm, though this was minimised with blocking during randomisation.

Participation

The population of interest were Centrelink customers who required an outbound contact from the Families, Child Care, Parenting, Students and Trainees program areas, and other general work that may have been allocated to the staff participating.

Inclusion Criteria

- a. Work has been allocated through the Workload Management System (WLM) to service officers with the appropriate skills in the selected teams and/or Smart Centre participating in the trial.
- b. Customer requires an outbound call, based on service officer's application of relevant processes.

Exclusion Criteria

Some exclusions were applied by service officers during the trial.

- a. Customers not registered for electronic messaging (the system cannot send SMS to customers not registered)
- b. Customer has opted out of Services Australia research
- c. Customers experiencing Family Domestic Violence if indicated on their record
- d. Level 1 and Level 2 Complaints and Feedback
- e. Customer Identified Risk of Homelessness
- f. Vulnerable Customer activities
- g. Internal reviews
- h. Urgent Payments

Other exclusions were applied by analysts when curating trial data.

- a. Service officers that do not log into softphone will not be captured.
- b. Dialed phone numbers that cannot be matched to a single customer. Outbound calls are not consistently associated with a particular customer, so the customer is identifiable if the dialed phone number can be matched exactly or with high confidence.
- c. Some phone numbers are truncated in the source data asset. A data matching process can be used to match some with high confidence, but not all will be included.
- d. Some SMS are sent to customers, though have not been matched them to any particular call or work item.
- e. All calls from service officers that are outside of the trial teams are out of scope.

Assumptions

- We assume customers receive an SMS at the time recorded in correspondence data. In some instances, there may be a network delay, though the extent to which this occurs is unknown.
- Service officers may send the Pre-Call SMS multiple times to the same customer. Multiple SMS from the same service officer to the same customer within one calendar day are considered part of a single overall outbound attempt; if any of the calls are successful then the outcome will be a one.
- Outbound contact processes generally require service officers to attempt all current phone numbers. Multiple calls from the same service officer to the same customer within one calendar day are considered part of a single overall outbound attempt; if any of the calls are successful then the outcome will be a one.

Statistical Modelling

The topic subject to a statistical test is the effect of the Pre-Call SMS on outbound contact success.

Point estimates, confidence intervals and p-values will be derived from an OLS regression model with the following specification:

$$Y_i = \beta_0 + \beta_t X_{it} + \beta_c Z_{ic} + \beta_{tc} X_{it} Z_{ic} + \epsilon_i$$

Where:

- Subscript i refers to the i th observation from the trial, from the first observation to the number of observations made. Like a simple average calculation, the model brings together all aspects across observations and gives a view of the trial topic according to the model specification.
- The dependent variable Y_i is the binary contact success outcome. For each observation, if $Y_i = 1$ the customer answers the phone and maintains connection for at least 30 seconds, or $Y_i = 0$ if not.
- β_0 is the intercept, which will measure success without the presence of the Pre-Call SMS. β_0 can also be written as $\beta_0 X_{i0}$. The success measure without an SMS can be calculated as $Y = \beta_0$.
- The X_{it} term is the treatment assignment indicator, where $\{0\}$ = control and $\{1\}$ = treatment 1/2. The β_1 coefficient represents the average treatment effect for the pooled interventions relative to control. The success measure can be calculated as $Y = \beta_0 + \beta_1$.
- The Z_{ic} terms refer to covariates expected to relate to both the outcome and the treatment. Covariates include stratification variables used when randomising into treatment and control groups. Covariates should be mean-centred where appropriate. Stratification covariates should have no statistical significance in the model.
- $X_{it} Z_{ic}$ represents the multiplication of X_{it} and Z_{ic} . When included in the regression model, the model generates parameter β_{tc} , which represents influence of covariates on treatment/control groups.
- ϵ_i is the cluster-robust standard error term.

OLS regression models provided point-estimates and standard errors required to perform the hypothesis tests.

The primary outcome used an intent-to-treat (ITT) framework and consisted of a covariate-adjusted comparison of the primary outcome.

Comparing SMS Types

Two conclusions were drawn by the broader group when designing the trial for which this report is the subject:

- The trial could not reasonably be powered to identify a difference of 1.2 percentage points between two SMS variants if it exists
- Even if it could, a 1.2 percentage point difference was not expected to be a meaningful difference to motivate permanently adjusting the Pre-Call SMS

Alongside the difference between SMS types, the standard error required to perform a statistical test was provided (see *Accessible Graph Data* in *Appendix A: Additional Calculations and Results*). Insufficient power aside, that statistical test would fail to reject the null hypothesis that there is no difference between SMS types. In other

words, there is no statistical evidence of a difference, even though the existing SMS appeared to slightly outperform the revised SMS.

The difference between SMS type effect on success was 2.5⁶ percentage points, higher than the expected 1.2 percentage point difference from the previous trial. One difference was the topic; this trial sent a Pre-Call SMS, while the previous trial sent a reminder SMS to report.

In considering whether a 2.5 percentage point difference is meaningful to us, it can be compared to the:

- expected 1.2 percentage points
- 12 percentage point effect between sending and not sending an SMS

The estimated difference is smaller than the difference between sending and not sending an SMS, so the general act of sending any SMS should be the key insight. However, a useful thought exercise may be why the revised SMS was less effective.

The report leads with the conclusion from this reasoning: *it is concluded there was no meaningful difference between the SMS types.*

Randomisation

Clustered randomisation occurred at the team level, where each cluster applies its assigned trial approach to outbound contacts it makes. After consideration of both service officer and team level randomisation, it was determined randomisation at the team level is the lowest level feasible, which is expected to dampen the benefit of what observations we capture (see *Power Analysis in Appendix D: Technical Details*).

'Work items', each of a particular 'work type', are assigned to service officers using a workload management program. 'Program Area', which takes either 'Families' or 'Students' values, was identified as a stratification variable for randomisation because prior knowledge demonstrates differences in SMS success across program areas. Service officers cannot control which work items are assigned to them.

Some teams or service officers do more relevant work than others. Stratification using a 'Work Volume' data point helped avoid one trial arm generating more relevant observations than another. To simplify stratification, we used the ordinality of the continuous Work Volume measure across participating teams to create 'Work Size', which takes either 'Small', 'Medium', or 'Large' values. The smallest 1/3 of the data (i.e. the teams producing the smallest volumes of relevant work) was assigned 'Small', the middle 1/3 'Medium' and the largest 1/3 'Large'. This was expected to account for cases where service officers would complete limited work of interest, such as working limited hours and/or performing non-target work. Exclusion criteria referred to the data capture issue of outbound contacts made while not logged in; outbound contacts were used as the 'Work Volume' data point, which may have handled this issue.

Randomisation was performed by establishing a list of teams to randomise. The list was imported into the *SAS Enterprise Guide* (SAS) statistical program. The list was processed through the *SURVEYSELECT* coding procedure, through which a pseudo-randomisation algorithm equally split teams to one of the three trial groups, also balancing program area and work volume across the groups. The list of teams and trial group assignments was exported and distributed to relevant team leaders for distribution with their teams. This randomisation was observed by three different branches within the agency and a representative of the Australian Centre for Evaluation (ACE).

During randomisation preparation, it was noted that only three Students teams were expected to be involved in the trial. Normally, two stratification variables would be specified as part of the *SURVEYSELECT* procedure. Where only three 'Students' teams appear available, the procedure would fail, given we require nine 'Students' teams: one team in each trial arm for each strata (unique combinations of stratification variables). The solution, as advised by ACE, was to create a hybrid stratification variable, including the values 'Small', 'Medium', 'Large' and 'Students'. All except the 'Students' group constitute the 'Families' teams. By effect, the procedure will randomise 'Families' teams stratified by the 'Work Size' values and simultaneously 'Students' teams without stratification.

⁶ Figure 1 presents 2.6 percentage points due to rounding error. The difference itself is -2.533 percentage points.

The trial relies on the integrity of randomisation. The standard is for all service officers randomised into a trial arm to participate. Some service officers may leave or enter; the extent to which this occurred and influenced results was discussed.

Power Analysis

We calculated the minimum number of observations required under different trial circumstances to inform which sites should be included in the trial. Power is the probability of detecting a difference when there is one and usually takes an academic default of 80%⁷.

A power analysis was performed using the [POWER](#) SAS procedure. The procedure used the following parameters:

Table 10: Parameters for Power Analysis

Power Analysis Parameter	Selection
Test Type	Proportion/mean difference
Statistical Test	Pearson's chi-square test / t-test
Significance Level	5% (statistical default)
Desired Power	80% (statistical default)
Test Sides	1 (superiority trial)
Group Proportions	1:2 (two groups; three arms with treatment pooling)
Null Proportion	0.56 (prior exploratory research)
Minimum Detectable Effect	0.06 (prior exploratory research)

These parameters suggested 6,000–8,000 observations required for a rule-of-thumb 80% power.

Clustering

Intra-cluster Correlation (ICC) captures the potential effect clustering may have on inference⁸. The reason why a relatively large number of observations (6,000–8,000) was suggested is because the smaller the number of clusters, the greater the potential effect they have on inference. Influential factors include:

- Team similarity, where more successful teams may be placed together in a trial arm, exaggerating or dampening a representative treatment effect
- Team uniqueness, where team members do not give a natural variety of data points, restricting use of standard errors to perform statistical tests

The academic expectation is that ICC sits between 0.01–0.02, giving the expectation of 57–77% power. This is short of the desired 80% power, though the minimum detectable effect is conservative for the Families cohort.

The final number of observations was 4,833. The ICC was 0.010854, on the lower side in the range of expectation. Largely due to the higher treatment effect (12%, rather than 6%), power was in excess of 99%, so we expect to be able to identify a statistically significant treatment effect if there was one.

⁷ Kemal (2020).

⁸ Killip, Mahfoud and Pearce (2004).

Statistical Inference

Statistical inference was performed using the [*SURVEYREG*](#) SAS procedure for OLS regression outputs and *SURVEYLOGISTIC* SAS procedure for logistic regression outputs.

Marginalisation

The effects of the Pre-Call SMS given in the report are calculated as ‘marginal average treatment effects’⁹. ‘Marginal’ can be read as ‘unconditional’, that is to say, an overall average treatment effect regardless of the subgroup. Stratification and blocking were performed during randomisation to accommodate imbalances in workload across teams, resulting in four subgroups: three Families subgroups and one Students subgroup. All models include this stratification, providing as many treatment effects ‘conditional’ on the observation coming from that group. ‘Marginalisation’ is the weighted average of treatment effects, creating an ‘unconditional’ average treatment effect not specific to any particular team in the trial. The same marginalisation logic was applied to baseline estimates and used for all summary statistics. As accessible results show, different subsets of the 4,833 observations affect the weighting, therefore the estimates.

⁹ Harrell (2015), University of Illinois (n.d.).

References

Hopewell S, Chan AW, Collins GS, Hróbjartsson A, Moher D, Schulz KF, et al. CONSORT 2025 Statement: updated guideline for reporting randomised trials. *BMJ*. 2025; 388:e081123. <https://www.bmj.com/content/389/bmj-2024-081123> Copyright 2025 Hopewell et al.

Brennan C Kahan, Joanna Hindley, Mark Edwards, Suzie Cro, Tim P Morris. The estimands framework: a primer on the ICH E9(R1) addendum. *BMJ*. Retrieved from <https://www.bmj.com/content/384/bmj-2023-076316>

Behavioural Economics Team of the Australian Government (BETA). *Effective use of SMS: timely reminders to report on time*. December, 2017. 978-1-925362-98-5 / 978-1-925362-99-2.

University of Illinois. *Measuring Treatment Effects*. Note: the original source is from the personal website of academic Jeoffrey D. Michler, currently an Associate Professor at the University of Arizona as of 2026. At the time of publishing, Jeoffrey's website redirects to another website that may be malicious, so the link has not been provided.

Hellevik, O. (2009). *Linear versus logistic regression when the dependent variable is a dichotomy*. *Quality & Quantity* 43(1):59-74. Retrieved from https://www.researchgate.net/publication/225208510_Linear_versus_logistic_regression_when_the_dependent_variable_is_a_dichotomy

McCoy, C. E. (2017). *Understanding the Intention-to-treat Principle in Randomized Controlled Trials*. *West J Emerg Med*. 2017 Sep 18;18(6):1075–1078. doi: 10.5811/westjem.2017.8.35985. Retrieved from <https://pmc.ncbi.nlm.nih.gov/articles/PMC5654877/>

Harrell, F. (<https://stats.stackexchange.com/users/4253/frank-harrell>), *The difference between average and marginal treatment effect*, URL (version: 2015-11-20): <https://stats.stackexchange.com/q/182787>. Note: while a statistical forum is not a reliable source, Frank Harrell is an acclaimed Professor of Biostatistics.

Kemal Ö. Power Analysis and Sample Size, When and Why? *Turk Arch Otorhinolaryngol*. 2020 Mar;58(1):3-4. doi: 10.5152/tao.2020.0330. Epub 2020 Mar 1. PMID: 32313887; PMCID: PMC7162597. Retrieved from <https://pmc.ncbi.nlm.nih.gov/articles/PMC7162597/>

Killip S, Mahfoud Z, Pearce K. What is an intracluster correlation coefficient? Crucial concepts for primary care researchers. *Ann Fam Med*. 2004 May–Jun;2(3):204–8. doi: 10.1370/afm.141. PMID: 15209195; PMCID: PMC1466680. Retrieved from <https://pmc.ncbi.nlm.nih.gov/articles/PMC1466680/#r2>

New South Wales Government (2022). *5 tips to maximise the impact of government text messages*. Last updated 07/07/2022. Retrieved from <https://www.nsw.gov.au/departments-and-agencies/behavioural-insights-unit/blog/maximise-text-message-impact>

Dal Grande, E., Chittleborough, C.R., Campostrini, S., Dollard, M. and Taylor, A.W., 2016. *Pre-survey text messages (SMS) improve participation rate in an Australian mobile telephone survey: an experimental study*. *PLoS One*, 11(2), p.e0150231. Retrieved from <https://pmc.ncbi.nlm.nih.gov/articles/PMC4769066/>

servicesaustralia.gov.au