Building a New Life in Australia:
The Longitudinal Study of Humanitarian Migrants

Data Users Guide

Release 5.0

June 2019

Australian Government
Department of Social Services
Australian Institute of Family Studies

MULTICULTURAL MARKETING & MANAGEMENT

colmar brunton.

The *Building a New Life in Australia* study was initiated and is funded by the Australian Government.

Contents

List of tables

List of figures

Abbreviations

AAPOR	American Association for Public Opinion Research
ABS	Australian Bureau of Statistics
ADA	Australian Data Archive
AIFS	Australian Institute of Family Studies
ANZSCO	Australian and New Zealand Standard Classification of Occupations
ASCED	Australian Standard Classification of Education
ASCL	Australian Standard Classification of Languages
ASCRG	Australian Standard Classification of Religious Groups
BNLA	Building a New Life in Australia
BVE	Bridging Visa Class E
CAPI	Computer-assisted personal interview
CASI	Computer-assisted self interview
CATI	Computer-assisted telephone interview
CBSR	Colmar Brunton Social Research
DIAC	Department of Immigration and Citizenship (now DIBP)
DIBP	Department of Immigration and Border Protection
DSS	Department of Social Services
ID	Identification number
MMM	Multicultural Marketing and Management
MU	Migrating Unit
NCLD	National Centre for Longitudinal Data
PAL	Primary Approach Letter
PA	Principal Applicant (see also PR)
PAPI	Pen and Paper Instrument
PaRMS	Participant and Respondent Management System
PR	Principal Respondent
PTSD	Posttraumatic Stress Disorder
SA	Secondary Applicant (see also SR)
SACC	Standard Australian Classification of Countries
SEIFA	Socio-Economic Indexes for Areas
SDB	Settlement Database
SHP	Special Humanitarian Program
SDQ	Strengths and Difficulties Questionnaire
SR	Secondary Respondent
UMA	Unauthorised Maritime Arrival
UNHCR	United Nations High Commissioner for Refugees

Key terms

Term	Meaning
Building a New Life in Australia (BNLA)	BNLA is a longitudinal study that aims to trace the settlement journey of humanitarian migrants, from their arrival in Australia through to eligibility for citizenship. It is also known as the *Longitudinal Study of Humanitarian Migrants*
Bridging Visa Class E (BVE)	A BVE is a temporary visa that allows people to stay in Australia while they finalise their immigration matter or make arrangements to leave Australia. BVE holders who arrived through the onshore migration pathway are not permitted to work in Australia unless they arrived prior to 13 August 2012.[1]
Community Detention[2]	Community Detention facilities house onshore pathway arrivals pending resolution of their claims for protection. During their stay they are not allowed to work, are subject to curfews and other supervision arrangements and can be moved back to detention at any time at the discretion of the Minister. Families, unaccompanied minors and other vulnerable clients are prioritised for placement in Community Detention.
Humanitarian Program	The Humanitarian Program is a component of Australia's Immigration Program and has two important functions: • The onshore protection/asylum component fulfils Australia's international obligations by offering protection to people already in Australia who are found to be refugees according to the United Nations Convention relating to the Status of Refugees. • The offshore resettlement component expresses Australia's commitment to refugee protection by going beyond these obligations and offering resettlement to people overseas for whom this is the most appropriate option.[3]
Migrating Unit (MU)	An MU comprises all persons who migrated to Australia as part of the same migration application as the Principal Applicant.

[1] Australian Human Rights Commission, *Tell Me About: Bridging Visas for Asylum Seekers*, accessed 6 March 2017 at https://www.humanrights.gov.au/our-work/asylum-seekers-and-refugees/publications/tell-me-about-bridging-visas-asylum-seekers
[2] The description of Community Detention provided here relates to the period relevant to the BNLA participants who may have spent time in such facilities. Policy changes since then and in the future may render this definition obsolete. For more information about the Community Detention policies in 2013 see: I Katz, G Doney & E Mitchell, *Evaluation of the expansion of the community detention program: client and service provider perspectives*, SPRC Report 12/13, University of New South Wales, 2013.
[3] DIBP, *Fact sheet - Australia's Refugee and Humanitarian programme*, accessed 24 February 2017 at https://www.border.gov.au/about/corporate/information/fact-sheets/60refugee

Term	Meaning

Offshore migration pathway	The offshore resettlement component of the Humanitarian Program comprises two categories of permanent visas. These are:
	• Refugee—for people who are subject to persecution in their home country, who are typically outside their home country, and are in need of resettlement. The majority of applicants who are considered under this category are identified and referred by the United Nations High Commissioner for Refugees (UNHCR) to Australia for resettlement. The Refugee category includes the visa subclasses of:
	  ○ Refugee (subclass 200)
	  ○ In-country Special Humanitarian (subclass 201)
	  ○ Emergency Rescue (subclass 203)
	  ○ Woman at Risk (subclass 204).
	• Special Humanitarian Program (SHP)—for people outside their home country who are subject to substantial discrimination amounting to gross violation of human rights in their home country, and immediate family of persons who have been granted protection in Australia. Applications for entry under the SHP must be supported by a proposer who is an Australian citizen, permanent resident or eligible New Zealand citizen, or an organisation that is based in Australia.[4] Successful applicants are granted a Global Special Humanitarian visa (subclass 202).[5]

Onshore migration pathway	The onshore component of the Humanitarian Program aims to provide options for people who wish to apply for protection (or asylum) after arrival in Australia.[6] It includes persons who either: • arrived without a valid visa (e.g. Unauthorised Maritime Arrivals (UMAs); or • were holders of valid visas (e.g. tourist) before claiming asylum. While onshore applications are being assessed, applicants may spend time in immigration detention, Community Detention or on a BVE. Australia currently has three types of visa subclasses available for people onshore who want to apply for protection: • Protection visa (subclass 866) • Temporary Protection visa (subclass 785) • Safe Haven Enterprise visa (subclass 790).[7] However, for the current BNLA survey participants, the only onshore migration visa that applied was subclass 866.

[4] DIBP, *Fact sheet - Australia's Refugee and Humanitarian programme*, accessed 24 February 2017 at https://www.border.gov.au/about/corporate/information/fact-sheets/60refugee

[5] DIBP, *Global Special Humanitarian visa (subclass 202)*, accessed 24 February 2017 at https://www.border.gov.au/Trav/Visa-1/202-

[6] DIBP, *Fact sheet - Australia's Refugee and Humanitarian programme*, accessed 24 February 2017 at https://www.border.gov.au/about/corporate/information/fact-sheets/60refugee

[7] DIBP, *Onshore – Protection*, accessed 1 March 2017 at https://www.border.gov.au/Trav/Refu/Onsh

Term	Meaning
Principal Applicant (PA)	A PA is the person on the visa application upon which the approval to immigrate was based. For the BNLA study, the PA is also designated as the lead participant for the household for Wave 1 of the study, if: • aged 18 years or older at the time of the first interview; and • had consented to participate in the study. A PA completes the BNLA PA survey questionnaire in Wave 1. (See also Principal Respondent.)
Principal Respondent (PR)	A PR is a BNLA study participant who is the lead participant for the household for a particular BNLA Wave, from Wave 2 onwards; and was either: • a PA; or • a Secondary Applicant who established a new household independent from the household in which their original PA lives; or • a Secondary Applicant who has been designated as the lead participant in a household for other reasons. A PR completes the BNLA PA survey questionnaire for a particular Wave from Wave 2 onwards.
Secondary Applicant (SA)	An SA is any member of the MU named on the visa application other than the PA (e.g. spouse or child of the PA). An SA is eligible to participate in the BNLA study if: • they were residing with their PA in Wave 1; and • their PA had consented to participate in the study in Wave 1. Furthermore, an SA for the BNLA study was either: • SA Adult—if aged 18 years or older at the time of the first interview and had consented to participate in the study; or • SA Adolescent—if aged 15 to 17 years at the time of the first interview, the adolescent's parent/guardian was a PA or an SA Adult for the study, and the parent/guardian had given consent at Wave 1 for the adolescent to participate in the study. SAs complete the BNLA SA survey questionnaire in Wave 1. (See also Secondary Respondent.)

Term	Meaning
Secondary Respondent (SR)	An SR is: • a BNLA study participant who is a member of a household where another person is the PR for the same BNLA Wave, from Wave 2 onwards; and • a respondent who completed the BNLA SA survey questionnaire for a particular Wave from Wave 2 onwards. Furthermore, an SR for the BNLA study could be either: • an Adult SR—if aged 18 years or over at the time of the interview and had consented to be interviewed for the study; or • an Adolescent SR—if aged 15 to 17 years at the time of the interview, the adolescent's parent/guardian was a PA or an SA Adult for the study, and the parent/guardian had given consent at Wave 1 for the adolescent to be interviewed. By wave 4 all Adolescent SR respondents had become Adult SR respondents.
Unauthorised Maritime Arrival (UMA)	A UMA is a person who arrives in Australia by boat without a valid visa.[8]

[8] For a full definition see Migration Act 1958, Section 5AA, at http://www.austlii.edu.au/au/legis/cth/consol_act/ma1958118/s5aa.html

Acknowledgements

A large number of people from a range of different organisations have had valuable input to the development and implementation of the Building a New Life in Australia study. It is not possible to list all of them here individually, but thanks are extended to staff and representatives from the following key organisations and forums:

- Department of Immigration and Border Protection (now known as the Department of Home Affairs)
- Department of Social Services—the National Centre for Longitudinal Data, and the Multicultural, Settlement Services and Communities Group (now known as the Settlement Services Group)
- Australian Institute of Family Studies
- Colmar Brunton Social Research
- Multicultural Marketing and Management
- The Building a New Life in Australia Survey Reference Group.

A range of individuals from other organisations also offered their time and advice on an ad hoc basis at various stages during development of the study. Their input is gratefully acknowledged here.

The hard work of the fieldwork team who were employed to locate study participants and undertake the household interviews is also acknowledged and we are grateful for all of their efforts.

Finally, we would like to extend a very special thank you to all of the humanitarian migrants who participated in the study. This study would not have been possible without their generosity in sharing their views and experiences.

1 Introduction

The Data Users Guide provides an overview of the Building a New Life in Australia (BNLA) project's background, aims and objectives and provides details of the study population, methodology, sample selection and characteristics, variable naming conventions and standardised scales used in the survey. It is intended that this Guide will be used in conjunction with the following documents:

- The Data Dictionary, which details the survey variables, administrative, derived and coded items to assist data users to analyse survey data.
- Survey questionnaires marked with variable names are also available on the study website and provide useful information for data users.

We invite feedback from data users about the BNLA data documentation. If there is anything you expected to find in the data users guide and didn't, if you need further clarification around any materials, or if you have any further comments, please let us know by emailing bnla@aifs.gov.au or NCLDresearch@dss.gov.au.

1.1 Background of the BNLA project

In 2012, the then Department of Immigration and Border Protection (DIBP), formerly named the Department of Immigration and Citizenship (DIAC), commissioned the BNLA project, known also as the *Longitudinal Study of Humanitarian Migrants*. The BNLA project aims to trace the settlement journey of humanitarian migrants, from their arrival in Australia through to eligibility for citizenship, in order to better understand the factors that influence people's settlement processes, both positively and negatively. BNLA is a ground-breaking project. There has been no project like this conducted in Australia before.

The Australian Institute of Family Studies (AIFS) has been commissioned to manage the design, administration and processing of the BNLA project. AIFS are partnering with Colmar Brunton Social Research (CBSR) in conjunction with Multicultural Marketing and Management (MMM), as the fieldwork agency for this project. In April 2014, responsibility for the project moved from the DIBP to the Department of Social Services (DSS) where it is now administered by the National Centre for Longitudinal Data (NCLD).

1.2 BNLA aims and objectives

The BNLA project is expected to provide a broad evidence base to assist policy development and program improvement for humanitarian migrants. In particular, it is expected to increase the knowledge base around factors that aid successful settlement and identify barriers that hinder positive outcomes. The project aims to ensure that future refugees experience a positive transition as they settle in Australia.

1.3 Research questions

A key interest for the research is to understand how different migration pathways affect settlement outcomes. The following broad research themes[9] have been formulated for this purpose:

- Humanitarian migrants' settlement outcomes in relation to their English language proficiency, housing situation, labour force participation, use of qualifications, income, health, community engagement, citizenship, and level of satisfaction with life in Australia.
- Access to and use of government and non-government services, welfare benefits and their effectiveness in contributing to migrants' successful settlement.
- How settlement outcomes of humanitarian migrants differ by visa subclass.

1.4 BNLA project overview

BNLA is a national project which focuses on how humanitarian migrants settle into life in Australia. Over its life, the project will follow the settlement journey of around 1,500 refugee migrating units and over 2,000 individuals within these migrating units. The study collects information from people who:

- received a permanent humanitarian visa overseas and arrived in Australia between May and December 2013 (i.e. the "offshore" cohort); or
- sought asylum after arriving in Australia and were subsequently granted a permanent humanitarian visa between May and December 2013 (i.e. the "onshore" cohort).

The project seeks to understand refugees' experiences with services, housing, education, employment, and their local community. The information gained will help to provide an overall picture of how Australia's new humanitarian migrants are settling and point to what can be done to improve the wellbeing of humanitarian migrants and their families. Information will be used to help researchers answer questions, such as:

- What is life like for a newly arrived humanitarian migrant?
- Are new humanitarian migrants healthy and happy?
- What difficulties are humanitarian migrants facing, and how are they coping?
- How well are services able to help humanitarian migrants settle?

Initial development work for the project commenced in September 2010 with the *Following Migrants Forward* workshop held at the Australian National University. Two Advisory Groups were then convened to inform the study design and development: a Survey Reference Group comprising Commonwealth government officials and external stakeholders and a Technical Advisory Group comprising experts in refugee resettlement and longitudinal survey methods and analysis. To further guide the project's development two background papers

[9] SE Khoo, *Key Research Questions for a Longitudinal Survey of Refugees and other Humanitarian Migrants*, Paper prepared for the Department of Immigration and Citizenship, Canberra, 2012.

were commissioned in 2012; one by Dr Siew-Ean Khoo[10] examining key issues and settlement indicators on which the survey should focus, and the other by Professor Matthew Gray and colleagues[11] outlining design options for the study. A position paper was also prepared by DIBP's Economic Analysis Unit, with input from the Survey Reference and Technical Advisory Groups.

Extensive scoping work was undertaken in 2012–13 during the development of the project. In-depth interviews and focus groups were conducted with more than 50 representatives from key stakeholders in the settlement sector including: Commonwealth and State Government departments; peak agencies that support refugees and asylum seekers; settlement service providers; migrant community groups and religious and cultural leaders; and former humanitarian migrants who had been in the community for a number of years. These consultations provided valuable advice around study design issues, participant recruitment and retention, survey content and cultural sensitivities in undertaking research with a vulnerable population.

[10] SE Khoo, *Key Research Questions for a Longitudinal Survey of Refugees and other Humanitarian Migrants*, Paper prepared for the Department of Immigration and Citizenship, Canberra, 2012.
[11] M Gray, A Graycar & L Nicolou, *Design options for the Building a New Life in Australia Longitudinal Survey of Humanitarian Migrants*, Paper prepared for the Department of Immigration and Citizenship, Canberra, 2012.

2 Project methodology

2.1 BNLA population and sample selection

The in-scope population comprised adult humanitarian migrants settling in Australia between May and October 2013 who had been issued with an appropriate permanent visa under the humanitarian entrant stream. The study aimed to include 1500 MUs (with 70% of participants being in offshore sub-groups and 30% in onshore sub-groups). This ratio broadly followed the distribution of government grant applications in the 2013 Humanitarian Program. To be eligible for the study, offshore visa holders had to have arrived in Australia holding a permanent visa three to six months prior to their Wave 1 interview. Onshore visa holders had to have received their permanent protection visa three to six months prior to their Wave 1 interview.[12] Given that Wave 1 interviewing was scheduled to take place from October 2013 to February 2014, the eligibility period for offshore arrivals and onshore visa grants was therefore May to December 2013.

With the implementation of offshore processing on 15 June 2013, and the transfer of UMAs from Australia to Nauru and Manus Island for processing, some changes to the sample composition were required. The revised sample parameters were an offshore component of 77% and an onshore component of 23%. Table 2.1 outlines the definitions and requirements for each cohort.

The migrating unit (MU) was the primary sampling unit, with the Principal Applicant (PA) selected as the main/lead participant for the project. The PA was identified in the Settlement Database[13] (SDB). The MU could consist of a single individual or members of a family.

The adult PA on the application was designated as the lead participant for the BNLA study (this was the person upon whom approval for a permanent visa was based). The PA (aged 18 years and over) was required to consent to take part in the study before other members of the family could be invited to participate.

A Secondary Applicant (SA) is any other member of the MU named on the visa application (e.g. spouse, children). SAs had to be aged 15 years and over and residing with the PA in order to participate in the study. However, once recruited to the study, SAs are able to independently continue participating in later waves if they choose (e.g. in circumstances where the PA withdraws or there has been a family break-up). In some cases SAs may establish a new household independent from the household in which their original PA continues to live. In such cases the SA will answer a PA questionnaire to ensure household information is captured.

[12] Onshore visa holders have been in Australia for a longer period, either on a different visa type or in immigration detention/Community Detention.
[13] The Department of Social Services is currently responsible for the Settlement Database.

Table 2.1: BNLA sample definitions and requirements

Cohorts	Permanent offshore (refugees)	Permanent onshore (asylum seekers)
Definition	People who came to Australia after being identified by the UNHCR as refugees in need of resettlement, plus people who came to Australia via the Special Humanitarian Program.	People who were already in Australia when they received a permanent visa.
Overall target	Sample target was *n* = 56 offshore MUs in the pilot and *n* = 1,160 offshore MUs in the main.	Sample target was *n* = 24 onshore MUs in the pilot and *n* = 340 onshore MUs in the main.
Sampling approach	A census approach was used in all selected locations given the relatively low numbers of eligible migrants potentially available.	A census approach was used in all selected locations given the relatively low numbers of eligible migrants potentially available.
Recruitment period	This group was recruited approximately 3 to 6 months after their arrival in the country.	This group was recruited approximately 3 to 6 months after being granted a permanent humanitarian visa rather than after their arrival in the country.

Subgroups	This group includes:
	• People under the Refugee Program
	  ○ Refugee subgroup, visa 200 (sample target = 940)
	  ○ Woman-at-risk, visa 204 (sample target = 170)
	  ○ Other, visa 201, 203 (sample target = 10).
	• People under the Special Humanitarian Program, visa 202 (sample target = 40).
		There are two distinct onshore sub-groups:
		• UMAs (those who arrived by boat without a valid visa) (sample target = 250)
		• Non-UMAs (those who applied for asylum after having arrived on a valid visa) (sample target = 90).

Fieldwork site selection	Fieldwork sites were selected on the basis that they would provide enough MUs for robust analysis of the women-at-risk and refugee subgroups and for the offshore cohort as a whole.	Fieldwork sites were selected on the basis that they would provide enough MUs for robust analysis of the UMA sub-population.

SAs are split into two sub-groups as follows.

- SA Adults: This sub-group comprises members of the PA's MU aged 18 years or older at the time of interview. There may be situations when members of the original MU were not living with the PA at the time of Wave 1 data collection. In such circumstances they were not eligible for the study. This could be the case for 'adopted families' where a group of unrelated migrants has formed an MU while en route to Australia. From Wave 2 onwards, we attempted to follow SAs who participated in Wave 1, but no longer live with the PA.
- SA Adolescents: These SAs are of particular interest and are eligible to be interviewed if their parent/guardian gave personal consent for participation at Wave 1. The eligible age for an SA Adolescent is 15 to 17 years. SAs younger than this cannot be interviewed. It is thought that adolescents will have different settlement outcomes from their parents or other adults in the MU and therefore it is important to capture their experiences.

2.1.1 Enumerated persons

PAs/PRs are asked to provide demographic information about other people who are living in the same household. The information includes age, sex and relationship to the PA/PR. In Waves 1 and 2, this information was collected only for the people who were on the original visa application (i.e. members of the same MU). Some of these people could be respondents in their own right (and complete the SA/SR questionnaire), but others only appear in the survey data as enumerated persons. Enumerated persons can be distinguished from respondents in the data using the *<a-e>resp*[14] variable.

From Wave 3 onwards, the collection of demographic information from the PR was expanded to include household members who were not on the original visa application and thus not part of the MU. For these additional enumerated persons, *xmu_id* (MU ID) values are assigned the value –9 "Not applicable". Also for those persons, the pre-arrival variables, such as *zapplicant* (Applicant type (Z: Pre-arrival))—but not the variables relating to pre-arrival trauma—are assigned the value –9 "Not applicable". In the Wave 3 dataset of BNLA Release 3, the variable *chhapp* (Applicant type (Wave 3)) was assigned the value 5 "New Member" for those persons. However, the *chhapp* has been dropped from subsequent releases of the Wave 3 dataset; it became apparent *chhapp* contained information that was not consistent with *zapplicant*.

Nonetheless, a New Member can be identified as these are the only observations where *zapplicant* is equal to –9 "Not applicable".

[14] See Section 3.2 for variable naming conventions.

2.2 Data collection

2.2.1 Wave 1 contact procedures

Information from the SDB was used to identify all permanent humanitarian visa holders residing in the selected sample locations who met the eligibility time frame described above[15]. Participants were eligible to be recruited from 11 sites around Australia. A total of 4,035 MUs were identified as potential participants in the study. Around half of the respective PAs were born in Iraq (26%) or Afghanistan (24%). Appendix A contains further information outlining the distribution of the sample in terms of country of birth.

The contact details of MUs comprising the eligible sample were supplied to CBSR. A letter of invitation to the study from AIFS was sent to the PA of each MU together with an accompanying information brochure. The letter and brochure were translated into the PA's primary language and contained the following information:

- a description of the study's rationale, aims, methodology, and topics to be covered in the interview
- an explanation of the differing roles of the organisations involved in the study
- reassurance that participation would be anonymous and the information participants provided would be confidential, could not be linked to a person's name and address and any identifying information would not be available to others
- an assurance that participation was voluntary and that acceptance or refusal of the invitation to participate (in Wave 1 or future waves) would not affect a person's visa status or access to services and supports
- details of the website established to provide further information about the study.

Following this initial contact, CBSR interviewers telephoned each potential participating individual/family to ascertain their interest in taking part in the study and to make an appointment if appropriate. If phone contact could not be made, interviewers undertook a home visit to try to reach potential study members. Community Engagement Officers were employed to advocate for the study, arrange introductions, and assist with participant recruitment in Wave 1. Their bilingual skills and familiarity with the local communities played a significant role in helping to locate MUs and build rapport with prospective participants.

For Wave 1, after all PAs who consented were interviewed, other members of the MU were invited to participate as SAs. The involvement of PAs and SAs is outlined further in Figure 2.1.

[15] More information about the SDB can be found at https://www.dss.gov.au/our-responsibilities/settlement-and-multicultural-affairs/programs-policy/settlement-services/settlement-reporting-facility

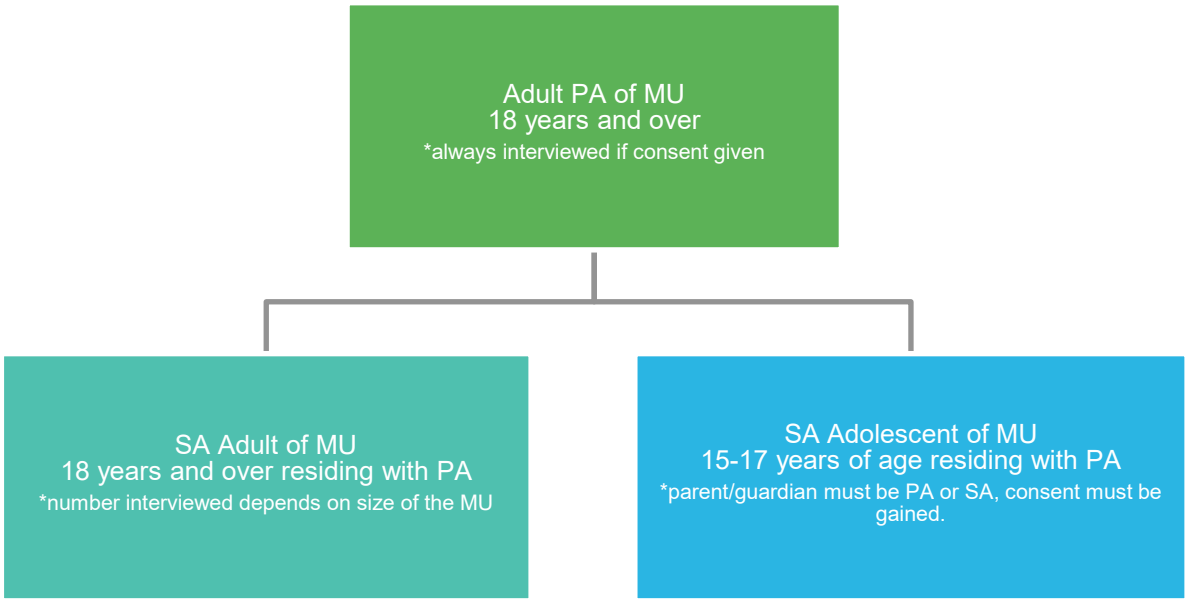

Figure 2.1: BNLA Wave 1 data collection with each MU

Starting from Wave 2, it is possible for SAs to exit the household of their original PA and establish their own household. In these cases, the exiting SA is asked to complete a household level questionnaire about their new household. As a result, from Wave 2 onwards, the migration status on arrival (PA or SA) may be different from the response status at the time of completing an interview. To this end, a person who completed a household level questionnaire is identified as a Principal Respondent (PR) while all other respondents are identified as a Secondary Respondent (SR).

2.2.2 Wave 2 contact procedures

In Wave 2, initial contact was through a Primary Approach Letter (PAL), which was addressed to each participant who had completed a Wave 1 interview. As participants' primary language was confirmed during Wave 1, the PAL used in Wave 2 was provided in this language. If participants' language was outside the set of languages translated for Wave 2, an English version was sent.

A number of strategies were put in place to increase the chances of making contact with participants in Wave 2, including:

- use of contact details collected from participants during the previous wave, such as mobile and home telephone number and email address
- use, if needed, of secondary contact details (of family or friends), provided by participants in the previous wave
- updated contact details supplied by DSS
- updated contact details provided by participants through the project website, 1800 hotline or email
- other methods of locating participants such as manual searches of White Pages or seeking assistance from other people in the household.

2.2.3 Subsequent waves

As part of the study design, the data collection method involved a shorter telephone interview in Wave 2, a face-to-face interview with a longer interview length in Wave 3, a shorter telephone interview in Wave 4, and the return of a face-to-face interview in Wave 5. Initial contact with the main sample was through a PAL, which was addressed to each eligible participant (i.e. separate letters to PAs, SA Adults and SA Adolescents) who had completed a Wave 1 interview and had not subsequently withdrawn from the study.[16] The PAL re-introduced the project and reminded participants that they would be contacted over the telephone by a field researcher to arrange a suitable time to complete the interview. The PAL contained the name of the field team member assigned to interview the respondent.

A study newsletter was also sent to all main and pilot participants at the start of each main phase. The aim of the newsletter was to share information gathered from the BNLA study with participants in order for them to see some of the key positive outcomes for recently arrived humanitarian migrants, hear further insights from the interviewing group and see some examples of how the data was being used by researchers and government.

As described in more detail later in Section 2.3.1, participants with children aged 5 to 17 years were eligible to participate in a new child module for the first time in Wave 3. There were two components associated with the child module. The first component consisted of the parent completing a child module. The second component of the child module involved a child self-report questionnaire. This child module was collected in Wave 3 only and not funded for inclusion in future waves of the study.

2.2.4 Interview mode

The initial five annual waves of data collection took place between 2013–14 and 2017–18, with alternating waves of home visits (Waves 1, 3 and 5) and telephone interviews (Waves 2 and 4). In Waves 1, 3 and 5, the survey was administered using one of the following three methods.

- A computer-assisted self-interview (CASI) on a small portable computer tablet. In Wave 1, the CASI had audio and flashlight functions available so that participants could listen to the questions and answers at the same time as the spoken words were highlighted on the screen. This feature was found to be particularly useful for participants with low levels of literacy[17].
- A computer-assisted personal interview (CAPI) was also offered to participants who preferred to complete the survey with an interviewer.
- When neither of these methods was feasible, participants were assisted by an accredited interpreter over the phone or in person, with an interviewer also present to pose the questions and record the answers.

[16] Participants who had withdrawn from the study after Wave 1 were not contacted.
[17] Due to low take-up rates, the audio and flashlight functions were not available from Wave 2 onwards.

An example of how questions are displayed on screen is provided in Figure 2.2 and Figure 2.3.

In Waves 2 and 4, the survey was administered over the telephone as a computer-assisted telephone interview (CATI) using bilingual interviewers and interpreters where required. However, in the final week of the Wave 2 fieldwork, in an attempt to boost response rates, a limited number of face-to-face interviews were conducted with respondents who were unable to be contacted through other methods.

In Waves 1, 3 and 5, the survey was administered as a face-to-face interview, either as a computer-assisted self-interview (CASI) or computer-assisted person interview (CAPI). However, there were some instances in which the interviewer was unable to schedule a face-to-face interview with the respondent due to the respondent's availability. In an attempt to reduce sample loss, a small number of CATI interviews were conducted with respondents who were unavailable for a face-to-face interview and/or requested a telephone interview during the fieldwork period in Wave 3 and Wave 5. The mode of survey completion for Waves 1, 3 and 5 are presented in Table 2.2.

A variable, <*a–e*>*mode*, in the dataset indicates the interview mode used for each participant.

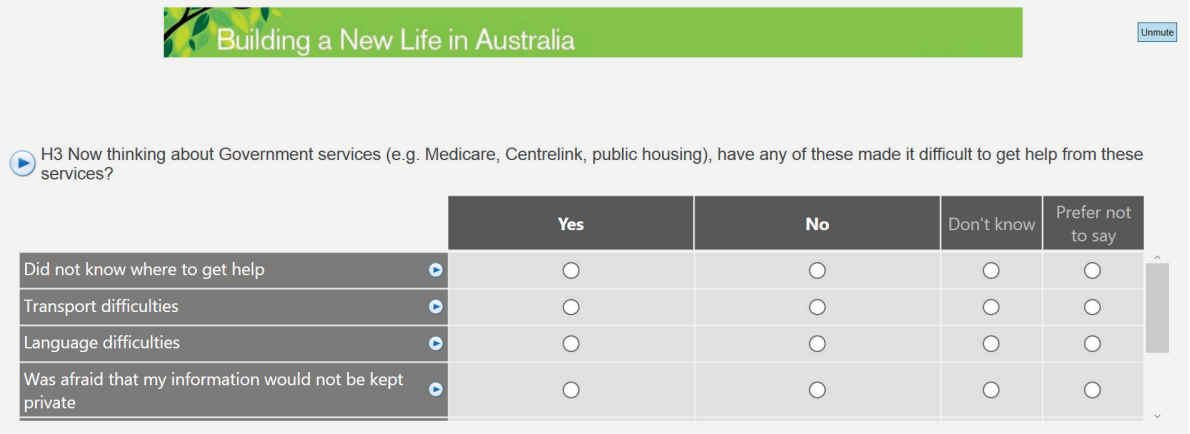

Figure 2.2: Screenshot of Question H3 from BNLA CASI

Figure 2.3: Screenshot of Question F22x from BNLA CASI

Table 2.2: Mode of survey completion for Waves 1, 3 and 5

Mode of interview	Wave 1		Wave 3		Wave 5	
	n	%	n	%	n	%
Computer Assisted Self-Interview (CASI)	1,692	70.5	1,999	63.3	847	45.0
Computer Assisted Personal Interview (CAPI)	707	29.5	640	33.8	941	50.0
Computer Assisted Telephone Interview (CATI)	0	0.0	55	2.9	93	4.9
Total interviews	2,399	100.0	1,894	100.0	1,881	100.0

2.2.5 Foreign language translations

To accommodate the diverse cultural and linguistic backgrounds of individuals in the study, the survey and participant materials were offered in 14 languages in Wave 1, covering approximately three-quarters of the languages spoken by the in-scope population. With the availability of interpreters, 19 languages were used across the Wave 1 sample. The top 10 languages used in Wave 1 by interview mode are listed in Table 2.3.

To facilitate comparison across the waves, Table 2.4 lists the main languages of survey completion for each of the five waves. For Wave 2, the questionnaire was translated into nine languages reflecting the major languages used in Wave 1. A total of nineteen languages were again used in Wave 2, while interviews were completed in 13, 11 and 10 different languages in Waves 3, 4 and 5, respectively.

Around 10% of respondents completed the survey in English in Waves 1 and 2. A slightly higher proportion of respondents completed an English survey in Wave 3 (12%), however the corresponding proportion was lower in Wave 4 (4%). At the time of the Wave 5 survey, close to one-quarter (23%) of respondents completed the survey in English.

Table 2.3: Wave 1 language of survey completion by interview mode

Interview language	CASI[(a)]	CAPI with interviewer[(b)]	CAPI with interpreter[(c)]	All modes
Arabic	43.9%	40.7%	8.2%	42.3%
Burmese/Myanmar	3.3%	2.3%	4.1%	3.0%
Chin Haka	1.8%	2.1%	np	1.9%
Dari	6.4%	12.3%	22.4%	8.3%
English	9.6%	10.2%	np	9.6%
Hazaragi	3.3%	3.6%	np	3.3%
Nepali	5.6%	np	10.2%	4.2%
Persian	22.6%	26.3%	12.2%	23.4%
Swahili	np	np	24.5%	1.2%
Tamil	1.3%	np	np	1.0%
Other[(d)]	2.2%	2.4%	18.4%	1.8%
Total	**100.0%**	**100.0%**	**100.0%**	**100.0%**
Number of interviews (*n*)	**1,692**	**658**	**49**	**2,399**

(a) The frequency for Swahili was less than 1% for CASI.
(b) The frequency for Nepali, Swahili, and Tamil were all individually less than 1% for CAPI with interviewer.
(c) The frequency for Chin Haka, English, Hazaragi, and Tamil were all individually less than 1% for CAPI with interpreter.
(d) Languages with less than 1% frequency are grouped together in the 'Other' category.
Note: Data are unweighted.
np Not published

Table 2.4: Waves 1-5 language of survey completion, all interview modes

Interview language	Wave 1	Wave 2[(a)]	Wave 3[(b)]	Wave 4[(c)]	Wave 5[(d)]
Arabic	42.3%	45.1%	44.2%	47.4%	40.8%
English	9.6%	9.1%	12.1%	4.3%	22.9%
Persian	23.4%	19.0%	24.5%	18.1%	19.3%
Dari	8.3%	10.8%	6.6%	8.3%	4.6%
Nepali	4.2%	2.5%	4.1%	5.4%	3.9%
Hazaragi	3.3%	5.0%	2.6%	9.4%	2.7%
Chin Haka	1.9%	2.7%	1.8%	3.2%	2.4%
Burmese/Myanmar	3.0%	2.5%	2.4%	2.4%	2.1%
Swahili	1.2%	np	np	np	np
Tamil	1.0%	1.6%	np	1.2%	np
Other[(e)]	1.8%	1.7%	1.7%	0.3%	1.3%
Total	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**
Number of interviews (*n*)	**2,399**	**2,009**	**1,894**	**1,929**	**1,881**

(a) For Wave 2, Swahili had a frequency of less than 1%.
(b) For Wave 3, Swahili and Tamil individually had a frequency of less than 1%.
(c) For Wave 4, Swahili had a frequency of less than 1%.
(d) For Wave 3, Swahili and Tamil individually had a frequency of less than 1%.
(e) Languages with less than 1% frequency for a Wave are grouped together in the 'Other' category.
Note: Data are unweighted.
np Not published

2.2.6 Fieldwork dates

Data are collected annually in two phases. A pilot study is conducted in May and June to test the study methodology and interview content. The main wave data collection takes place between October and the following February or March. Table 2.5 displays the periods that each of the completed BNLA fieldwork activities took place.

Table 2.5: BNLA fieldwork time periods for pilot studies and main waves

Wave	Pilot study	Main wave
1	May & June 2013	Oct 2013 to Mar 2014
2	May & June 2014	Oct 2014 to Feb 2015
3	May & June 2015	Oct 2015 to Feb 2016
4	May & June 2016	Oct 2016 to Feb 2017
5	May & June 2017	Oct 2017 to Mar 2018

After Wave 1, a minimum of 9 months was required to elapse between each main wave interview. For example, the earliest time a BNLA respondent was interviewed for Wave 1 was 3 months after arrival in Australia (if in the offshore cohort) or 3 months after a permanent visa was granted (if in the onshore cohort). For Wave 2, for the offshore group the earliest was 12 months after arrival, and for the onshore group the earliest was 12 months after being granted a permanent visa.

2.2.7 Interview length

PRs complete a longer questionnaire than SRs as they answer additional questions about their family and housing that is only collected once for each household (including but not limited to household demographics). Table 2.6 details average interview length for different respondent types in each wave.

In Wave 1, some variation in survey timings between subgroups was found, with the onshore PA group taking less time on average compared to offshore PA. Interviews completed with a bilingual interviewer were quicker on average compared to self-completed interviews where the respondent completed the interview on the computer tablet.

In Wave 2, a considerably reduced survey instrument was used to accommodate a shorter target interview length for administration over the telephone. There was little variation by language for interviews administered over the telephone, however the small number of surveys conducted face-to-face as CAPI interviews were much shorter on average (15 minutes for PRs and 7 minutes for SR Adults), highlighting that interviews conducted in face-to-face mode were quicker to administer than those over the telephone.

As part of the study design, the data collection method in Wave 3 reverted back to a face-to-face interview and a longer interview length. There were differences by interview mode across all applicant types, with CAPI interviews taking longer on average than CASI interviews. More specifically, CASI interviews in English took 45 minutes on average to complete compared to CAPI interviews in English, which took just over 1 hour to complete (for PRs).

In Wave 4, data was again collected via telephone interview. CATI interviews in English took longer on average than CATI interviews in a translated language among PRs. Given that only 5% of PRs completed an interview in English, this greater average survey length did not have a significant impact on the overall survey timings. There were no differences in survey length among SRs by interview mode.

Table 2.6: Average BNLA interview length, Waves 1 to 5

Wave	PA/PR	SA/SR Adult	SA/SR Adolescent
1	56 minutes	40 minutes	35 minutes
2	23 minutes	11 minutes	9 minutes
3	50 minutes	30 minutes	26 minutes
4	20 minutes	11 minutes	N/A
5	48 minutes	34 minutes	N/A

In Wave 5, the data collection method again reverted back to a face-to-face interview. With the addition of new survey material, including a new module to capture respondents' attitudes towards childcare and gender roles, the target survey length was reduced by 5 minutes for PRs and increased by 5 minutes for SRs to ensure the overall combined target of 90 minutes was achieved. In contrast to previous waves, CASI interviews in a language other than English took longer on average than other modes of administration (55 minutes for PRs and 43 minutes for SRs.

2.3 Survey instruments

The BNLA study collects data on a wide range of topics. The survey content was developed based on extensive consultations undertaken by AIFS prior to the commencement of Wave 1. These consultations and associated scoping work involved key stakeholders in the settlement sector, such as representatives of peak agencies, settlement service providers, community, cultural and faith-based groups, and humanitarian migrant communities. The main focus of these consultations was to seek advice about the study design and methodology, recruitment and retention strategies, survey content, and cultural sensitivities in undertaking research with this population.

The major domains of the questionnaire are detailed in Table 2.7 and include:

- family composition and demographics
- housing and neighbourhood characteristics
- English language proficiency and training
- engagement in other education, study or training
- employment, income and financial hardship
- migration pathways and experiences
- physical and mental health
- self-sufficiency
- community engagement and support
- perceptions of life in Australia
- child module (which was only used in Wave 3)
- childcare and gender roles (new module introduced in Wave 5).

Table 2.7: BNLA topics

Module Code	Topic	Scope
de	Family composition and demographics	Demographic information relating to the family such as age, gender of family members, country of birth and marital status
ho	Housing and neighbourhood	Assistance in finding housing, number of times moved home, tenure type, quality of housing, number of bedrooms, neighbourhood characteristics
lp	English language proficiency	Languages spoken at home, English language proficiency, whether attending English language classes, use and helpfulness of interpreting services
ed	Education and training	Highest level of education achieved, current education and training undertaken, educational aspirations, previous qualifications gained prior to arrival in Australia and whether they have been recognised
em	Employment and income	Current employment status, employment characteristics, prior occupation and work experience, experience of unemployment in Australia, income and government benefits received, financial stress
ie	Immigration experience	Life before settling in Australia including countries resided in prior to arrival, the experience of deprivation or trauma, time spent and type of services accessed in refugee camps, Australian immigration detention or Community Detention, reasons for migrating to Australia, social networks available upon arrival
he	Health	Information on physical and mental health, life stressors and coping
ss	Self-sufficiency	Experiences and ease of accessing services, information and transport, barriers to service use
cs	Community support	Levels of support from national, religious and other community groups, involvement in community activities, ease of making friends, sense of belonging in Australia
ls	Personal resources and life satisfaction	Satisfaction with current life and situation, self-concept, self-efficacy, levels of trust in different community groups and organisations, experience of discrimination
la	Life in Australia	Expectations of life in Australia before arrival, factors promoting or hindering settlement, ease of settling in Australia
cm	Child module	Intergenerational transmission of trauma in migrant communities (Parent report)
sr	Child module	Intergenerational transmission of trauma in migrant communities (child self-report)
cg	Childcare and gender role attitudes	Caregiver status, childcare arrangements, attitudes toward gender roles and work

The broad range of topics used in Wave 1 was carried over to the Wave 2 survey instrument. As noted above, a reduction to the survey content was required to accommodate the shorter interview length in Wave 2. To this end, material only needed to be collected once was not re-asked in Wave 2, such as former countries of residence and experiences prior to arrival in Australia. A small number of new items were included in the Wave 2 survey to capture experiences relevant to later stages of settlement, such as feelings of connectedness to home language and culture, experiences of bullying (for adolescent respondents only), and additional demographic detail in cases where an SA has moved out of the original Wave 1 household. Other minor revisions to code-frames were made based on an analysis of their use in Wave 1. In Wave 3, respondents were re-asked about their pre-arrival experience of trauma. However the questions were asked in respect of personal experience only rather than that of themselves or other family members as asked in Wave 1. The exact list of traumatic events asked about was also different. In other respects, the Wave 3 instrument closely resembled the Wave 1 instrument, while the survey instrument used in Wave 4 was similar to the one used in Wave 2. At Wave 3, a new data collection module (C*hild Module*) focused on children was introduced. This aspect of the study is described further in the next section. With the exception of this child module, the majority of items used in the Wave 3 instrument were carried forward for use in the Wave 5 instrument. New material was added to the Wave 5 survey including, additional items to capture information about respondents who were unable to work or study due to family/childcare reasons, additional physical health items, and follow-up questions about experiences of discrimination. These were asked of both PRs and SRs. A new module (childcare and gender roles) was also introduced in Wave 5, capturing information relating to use of childcare and attitudes toward gender roles. The Data Dictionary provides a more detailed description of the survey items and changes implemented since Wave 1.

Survey instruments are supplied with the data as part of data release and are available on the BNLA study website at the following AIFS website address: http://www3.aifs.gov.au/bnla/

2.3.1 Wave 3 Child module

The Wave 3 survey content was almost identical to that used for Wave 1. The notable difference was the inclusion of a module focussed on children in Wave 3, a child module. Phoenix Australia contracted AIFS to develop and implement a child module for the BNLA study for Wave 3. (Phoenix Australia, which is also known as the Centre for Posttraumatic Mental Health, is a not-for-profit organisation affiliated with the Department of Psychiatry at the University of Melbourne.) The child module aimed to explore intergenerational transmission of trauma in migrant communities.

There were two components associated with the child module. The first component consisted of the parent completing a child module where primary caregivers (in most cases, mothers) were invited to answer questions for up to two of their children aged 5 to 17 years. The target length of each set of questions was 10 minutes of interviewing time per child. The child module survey questions were programmed into the PR and SR surveys and questions covered how the children were doing at school, whether they were making friends, their health and wellbeing and how they were adjusting to life in Australia. The Strengths and Difficulties

Questionnaire (SDQ)[18]—a brief behavioural screening questionnaire—was also included as part of this set of questions. This instrument is designed for administration with parents of children aged 4 to 17 years.[19]

The second component of the child module involved a child self-report questionnaire, which consisted of children aged 11 to 17 years completing a short Pen and Paper Instrument (PAPI). Up to two children per household were chosen to participate. Interviewers first obtained parental consent before the child was invited to complete the self-report questionnaire. The self-report questionnaire was completed in nine languages other than English. A major part of the questionnaire comprised the self-report SDQ for children aged 11 to 17 years.[20] There was some cross-over in the type of content asked in each questionnaire, with the SDQ, health status and trauma covered in both the parent complete and child complete questionnaire, to allow comparison of key outcomes for young people from the perspective of parents and the children themselves. Children aged 11 to 17 years were also asked to complete the PTSD-8 post-trauma inventory[21]; and answer questions covering antisocial behaviour, physical activity, and academic and sporting achievements and awards received.

Sample selection for the child module was undertaken by first randomly selecting two children aged 11 to 17 years in each household. These two randomly selected children were eligible for participation in the child self-report questionnaire sub-study, and parents were invited to answer child module questions for the same two children. Where there was only one child aged 11 to 17 years in the household, if there were other children aged 5 to 10 years then one was randomly selected to be eligible as the second child for the parent complete module. Where there were only children aged 5 to 10 years in the household, up to two children aged 5 to 10 years were randomly selected for participation in the parent complete module.

[18] © Robert Goodman
[19] youthinmind, *One-sided SDQ for parents or teachers of 4-17 year olds,* retrieved 18 April 2017 from http://www.sdqinfo.org/py/sdqinfo/b3.py?language=Englishqz(UK)
[20] youthinmind, *One-sided self-rated SDQ for 11-17 year olds,* retrieved 1 March 2017 from http://www.sdqinfo.org/py/sdqinfo/b3.py?language=Englishqz(Austral)
[21] M Hansen, TE Anderson, C Armour, A Elklit, S Palic & T Mackrill, 'PTSD-8: A Short PTSD Inventory', *Clinical Practice & Epidemiology in Mental Health*, vol. 6, 2010, pp. 101-108.

2.4 Fieldwork response rates

In Wave 1 a total of 1,509 PAs, 755 SA Adults and 135 SA Adolescents were interviewed, taking the total number of face-to-face interviews to 2,399. Appendix B provides further analysis of survey completion by visa subclass and gender.

Participation and response rates of PAs for Wave 1 of the study are detailed in Table 2.8. The table reports response rates in several ways in accordance with the standardised response rates developed by the American Association for Public Opinion Research (AAPOR). Firstly, response rates were examined as the proportion of the total eligible sample (n = 4,035). A rate of 37% was achieved using this criterion. However, this includes individuals/families who were not initiated for contact, could not be contacted, or who refused. Secondly, response rates can be examined as the proportion of the cohort who were initiated for contact (n = 2,769). Using this criterion, a rate of 55% was achieved (See Response Rate 2). Thirdly, using the criterion of the proportion who were successfully contacted (n = 2,031), a response rate of 74% was achieved (See Cooperation Rate 1). This includes people who were successfully contacted but could not proceed with an interview for a variety of reasons, such as the quota was met, they moved to an area outside of the scope of interviewing or were unavailable for the duration of the fieldwork period. Finally, a response rate was calculated using the proportion who were successfully contacted and eligible to proceed with an interview (Cooperation Rate 2). Using this criterion, Table 2.8 shows that 83% of respondents who were contacted and able to proceed with an interview agreed to participate, while 17% refused. Rates were very high across the visa subclasses (with the exception of the 866 non-UMA visa subclass group with a rate of 58%). Thus, inability to contact potential participants was the largest reason for non-response, while the rate of refusal was relatively low.

Table 2.7: PAs – Participation and response rates by visa subclass, Wave 1

Visa subclass	Initiated for contact	Completed interviews	Response rate 1	Response rate 2	Cooperation rate 1	Cooperation rate 2
200	1,589	942	41.6%	59.3%	80.4%	85.2%
201	19	8	33.3%	42.1%	66.7%	66.7%
202	79	42	37.5%	53.2%	70.0%	75.0%
204	292	183	45.6%	62.7%	83.6%	89.3%
866 UMA	453	234	42.2%	51.7%	70.5%	85.4%
866 non-UMA	336	100	14.8%	29.8%	42.2%	58.1%
Total	**2,769**	**1,509**	**37.4%**	**54.5%**	**74.3%**	**82.7%**

Notes: Formulas for calculating standardized response rates have been developed by the AAPOR: http://www.aapor.org/Education-Resources/For-Researchers/Poll-Survey-FAQ/Response-Rates-An-Overview.aspx

- Response rate 1 calculated as the proportion of the eligible sample who completed an interview.
- Response rate 2 calculated as the proportion of the sample initiated for contact who completed an interview. This is consistent with the way the AAPOR Response Rate 1 is calculated.
- Cooperation rate 1 calculated as the proportion of the sample successfully contacted who completed an interview. This is consistent with the way the AAPOR Cooperation Rate 1 is calculated.
- Cooperation rate 2 calculated as the proportion of the sample successfully contacted and eligible to proceed, who completed an interview. This is consistent with the way the AAPOR Cooperation Rate 3 is calculated.

It is important to note that, from Wave 2 onwards:

- SA Adolescents who had turned 18 answered the adult questions on the SA survey
- some original SAs completed a PR survey in cases where they had moved out of the original household and became the PR for the survey.

Therefore, the number of interviews by 'Survey Type' (the *<a-e>survey* variable) reflects the number of principal/secondary questionnaires completed, and from Wave 2 onwards, differs from the number of interviews by the original 'Applicant Type' (the *zapplicant* variable). In this section of the Data Users Guide, the numbers of interviews for each wave and calculation of retention, response and participation rates are all based on the original applicant type at Wave 1 (e.g. PA, SA Adult or SA Adolescent), regardless of the type of survey completed. This allows accurate tracking of retention for the different participant types (otherwise the number of completed SA Adult surveys would artificially increase because of the inclusion of SA Adolescents and there would be a corresponding apparent drop in the number of SA Adolescent surveys).

A total of 2,009 interviews were completed in Wave 2, comprised of 1,268 PAs, 626 SA Adults and 115 SA Adolescents. Appendix B provides further analysis of survey completion by visa subclass and gender.

Participation rates for Wave 2 are indicated in Table 2.9 for PAs and Table 2.10 for SAs. Two types of response rates were used to describe Wave 2 outcomes: Response Rate 2; and Cooperation Rate 2. The first response rate examines completed interviews as a proportion of the sample initiated for contact (PA: n = 1,509; SA: n = 890). Using this criterion, a rate of 84% for PAs and 83% for SAs was achieved (see Response Rate 2). A cooperation rate was calculated as the proportion who were successfully contacted and eligible to proceed with an interview who took park. Using this criterion, 95% of PAs and 96% of SAs who were successfully contacted agreed to take part.

There were some differences by visa subgroup. PAs in the 204 visa category (i.e., woman at risk) had the lowest response rate (77%) but one of the highest cooperation rates (97%). This indicates that once respondents in this subgroup could be contacted, almost all were happy to take part. Conversely, SAs in this category had higher response rates (84%). Those in the 202 visa category had lower rates compared to the other visa subgroups (83% among PAs and 70% among SAs). It can also be observed that PAs in the onshore group had the lowest co-operation rates of all of the visa subclasses (92% for the 866 UMA; 89% for the 866 non-UMA).

Table 2.8: PAs – Participation and response rates by visa subclass, Wave 2

Visa subclass	Eligible sample	Sample initiated	Successfully contacted & eligible to proceed	Completed interviews	Response rate 2	Cooperation rate 2
200	942	942	844	807	85.7%	95.6%
201	8	8	7	7	87.5%	100.0%
202	42	42	37	35	83.3%	94.6%
204	183	183	146	141	77.0%	96.6%
866 UMA	234	234	213	196	83.8%	92.0%
866 non-UMA	100	100	92	82	82.0%	89.1%
Total	**1,509**	**1,509**	**1,339**	**1,268**	**84.0%**	**94.7%**

Note: Formulas for calculating standardized response rates have been developed by the AAPOR: http://www.aapor.org/Education-Resources/For-Researchers/Poll-Survey-FAQ/Response-Rates-An-Overview.aspx
Response rate 2 calculated as the proportion of the sample initiated for contact who completed an interview. This is consistent with the way the AAPOR Response Rate 1 is calculated.
Cooperation rate 2 calculated as the proportion of the sample successfully contacted and eligible to proceed, who completed an interview. This is consistent with the way the AAPOR Cooperation Rate 3 is calculated.

Table 2.9: SAs – Participation and response rates by visa subclass, Wave 2

Visa subclass	Eligible sample	Sample initiated	Successfully contacted & eligible to proceed	Completed interviews	Response rate 2	Cooperation rate 2
200	686	686	604	580	84.5%	96.0%
201	2	2	2	2	100.0%	100.0%
202	47	47	38	33	70.2%	86.8%
204	109	109	93	92	84.4%	98.9%
866 UMA	26	26	24	21	80.8%	87.5%
866 non-UMA	20	20	14	13	65.0%	92.9%
Total	**890**	**890**	**775**	**741**	**83.3%**	**95.6%**

Note: Formulas for calculating standardized response rates have been developed by the AAPOR: http://www.aapor.org/Education-Resources/For-Researchers/Poll-Survey-FAQ/Response-Rates-An-Overview.aspx
Response rate 2 calculated as the proportion of the sample initiated for contact who completed an interview. This is consistent with the way the AAPOR Response Rate 1 is calculated.
Cooperation rate 2 calculated as the proportion of the sample successfully contacted and eligible to proceed, who completed an interview. This is consistent with the way the AAPOR Cooperation Rate 3 is calculated.

A total of 1,894 interviews were completed in Wave 3, which included 1,155 PAs, 624 SA Adults and 115 SA Adolescents. Participation rates for Wave 3 are indicated in Table 2.11 for PAs and Table 2.12 for SAs. The participation and response rates for Wave 3 follow the same approach as those calculated for Wave 2.

Firstly, looking at the proportion of the eligible sample at the beginning of the Wave 3 main fieldwork who completed a survey, an overall rate of 77% for PAs and 83% for SAs was achieved. Secondly, response rates were examined as the proportion of the sample who were successfully contacted and eligible to proceed. Using this criterion, a cooperation rate of 94% for PAs and 95% for SAs was achieved.

The cooperation rate data at Wave 3 also showed that the study continues to experience relatively low rates of refusal. However, there were some differences in participation rates by visa subgroups. For PAs, the 866 UMA group had the lowest overall response rate of 63% but a higher cooperation rate of 92% of eligible PAs.

Table 2.10: PAs – Participation and response rates by visa subclass, Wave 3

Visa subclass	Eligible sample	Sample initiated	Successfully contacted & eligible to proceed	Completed interviews	Response rate 2	Cooperation rate 2
200	939	939	790	752	80.1%	95.2%
201	8	8	7	6	75.0%	85.7%
202	41	41	29	28	68.3%	96.6%
204	183	183	150	147	80.3%	98.0%
866 UMA	98	98	174	62	63.3%	92.0%
866 non-UMA	229	229	82	160	69.9%	75.6%
Total	**1,498**	**1,498**	**1,232**	**1,155**	**77.1%**	**93.8%**

Note: Formulas for calculating standardized response rates have been developed by the AAPOR: http://www.aapor.org/Education-Resources/For-Researchers/Poll-Survey-FAQ/Response-Rates-An-Overview.aspx
Response rate 2 calculated as the proportion of the sample initiated for contact who completed an interview. This is consistent with the way the AAPOR Response Rate 1 is calculated.
Cooperation rate 2 calculated as the proportion of the sample successfully contacted and eligible to proceed, who completed an interview. This is consistent with the way the AAPOR Cooperation Rate 3 is calculated.

Table 2.11: SAs – Participation and response rates by visa subclass, Wave 3

Visa subclass	Eligible sample	Sample initiated	Successfully contacted & eligible to proceed	Completed interviews	Response rate 2	Cooperation rate 2
200	684	684	612	586	85.7%	95.8%
201	2	2	2	1	50.0%	50.0%
202	46	46	34	33	71.7%	97.1%
204	109	109	94	94	86.2%	100.0%
866 UMA	19	19	19	11	57.9%	73.7%
866 non-UMA	26	26	15	14	53.8%	73.3%
Total	**886**	**886**	**776**	**739**	**83.4%**	**95.2%**

Note: Formulas for calculating standardized response rates have been developed by the AAPOR: http://www.aapor.org/Education-Resources/For-Researchers/Poll-Survey-FAQ/Response-Rates-An-Overview.aspx
Response rate 2 calculated as the proportion of the sample initiated for contact who completed an interview. This is consistent with the way the AAPOR Response Rate 1 is calculated.
Cooperation rate 2 calculated as the proportion of the sample successfully contacted and eligible to proceed, who completed an interview. This is consistent with the way the AAPOR Cooperation Rate 3 is calculated.

A total of 1,929 interviews were completed in Wave 4, which included 1,186 PAs, 628 SA Adults and 115 SA Adolescents (although by this time all SA Adolescents were of an age to complete an adult interview, and there were no adolescent questionnaires administered in Wave 4). Participation rates for Wave 4 are indicated in Table 2.13 for PAs and Table 2.14 for SAs. The participation and response rates for Wave 4 follow the same approach as those calculated in pervious waves.

Based on the eligible sample in Wave 4, an overall response rate of 80% for PAs and 85% for SAs was achieved. These response rates were higher than those achieved in Wave 3 for both groups. The proportion of participants successfully contacted who completed an interview – the cooperation rate – was 96% for both PAs and SAs. This proportion was also higher than in Wave 3 for both respondent types. As in previous waves, there were differences in participation rates by visa subclass. PAs and SAs on 866 and 202 visas were more difficult to contact, as shown by the lower response rates. Nevertheless, participants on 202 and 866 UMA visas showed high cooperation rates. PAs and SAs on 866 non-UMA visas had the lowest cooperation rates in Wave 4 (84% and 80% respectively), although these cooperation rates were higher than in Wave 3.

Table 2.12: PAs – Participation and response rates by visa subclass, Wave 4

Visa subclass	Eligible sample	Sample initiated	Successfully contacted & eligible to proceed	Completed interviews	Response rate 2	Cooperation rate 2
200	928	928	806	784	84.5%	97.3%
201	8	8	7	7	87.5%	100.0%
202	40	40	33	29	72.5%	87.9%
204	182	182	156	154	84.6%	98.7%
866 UMA	224	224	151	146	65.2%	96.7%
866 non-UMA	93	93	79	66	71.0%	83.5%
Total	**1,475**	**1,475**	**1,221**	**1,186**	**80.4%**	**96.3%**

Note: Formulas for calculating standardized response rates have been developed by the AAPOR: http://www.aapor.org/Education-Resources/For-Researchers/Poll-Survey-FAQ/Response-Rates-An-Overview.aspx
Response rate 2 calculated as the proportion of the sample initiated for contact who completed an interview. This is consistent with the way the AAPOR Response Rate 1 is calculated.
Cooperation rate 2 calculated as the proportion of the sample successfully contacted and eligible to proceed, who completed an interview. This is consistent with the way the AAPOR Cooperation Rate 3 is calculated.

Table 2.13: SAs – Participation and response rates by visa subclass, Wave 4

Visa subclass	Eligible sample	Sample initiated	Successfully contacted & eligible to proceed	Completed interviews	Response rate 2	Cooperation rate 2
200	676	676	609	587	86.8%	96.4%
201	2	2	2	2	100.0%	100.0%
202	45	45	32	31	68.9%	96.9%
204	109	109	97	94	86.2%	96.9%
866 UMA	23	23	18	17	73.9%	94.4%
866 non-UMA	19	19	15	12	63.2%	80.0%
Total	**874**	**874**	**773**	**743**	**85.0%**	**96.1%**

Note: Formulas for calculating standardized response rates have been developed by the AAPOR: http://www.aapor.org/Education-Resources/For-Researchers/Poll-Survey-FAQ/Response-Rates-An-Overview.aspx
Response rate 2 calculated as the proportion of the sample initiated for contact who completed an interview. This is consistent with the way the AAPOR Response Rate 1 is calculated.
Cooperation rate 2 calculated as the proportion of the sample successfully contacted and eligible to proceed, who completed an interview. This is consistent with the way the AAPOR Cooperation Rate 3 is calculated.

A total of 1,881 interviews were completed in Wave 5, which included 1,144 PAs, 624 SA Adults and 116 SA Adolescents (although there were no adolescent questionnaires administered in Wave 5). Participation rates for Wave 5 are indicated in Table 2.15 for PAs and Table 2.16 for SAs. The participation and response rates for Wave 5 follow the same approach as those calculated in previous waves.

Based on the eligible sample in Wave 5 an overall response rate of 78% for PAs and 85% for SAs was achieved. These response rates were slightly higher than those achieved in Wave 3 for both groups. This higher response rate observed for SAs reflects that there were fewer challenges in locating and making contact with SA participants in Wave 5.

The proportion of participants successfully contacted who completed an interview – the cooperation rate – was around 95% for both PAs and SAs. As in previous waves, there were differences in participation rates by visa subclass, with the lowest response rates and cooperation rates observed among respondents holding an 866 non-UMA visa. The data presented in Tables 2.14 and 2.15 indicate that respondents in the 866 non-UMA subgroup were the most difficult to locate as well as being the least likely to take part in a Wave 5 interview.

Table 2.14: PAs – Participation and response rates by visa subclass, Wave 5

Visa subclass	Eligible sample	Sample initiated	Successfully contacted & eligible to proceed	Completed interviews	Response rate 2	Cooperation rate 2
200	922	922	793	753	81.7%	94.9%
201	8	8	7	7	87.5%	100.0%
202	41	41	33	32	78.0%	97.0%
204	181	181	153	149	82.3%	97.4%
866 UMA	223	223	159	150	67.3%	94.3%
866 non-UMA	91	91	70	53	58.2%	75.7%
Total	**1,466**	**1,466**	**1,215**	**1,144**	**78.0%**	**94.0%**

Note: Formulas for calculating standardized response rates have been developed by the AAPOR: http://www.aapor.org/Education-Resources/For-Researchers/Poll-Survey-FAQ/Response-Rates-An-Overview.aspx
Response rate 2 calculated as the proportion of the sample initiated for contact who completed an interview. This is consistent with the way the AAPOR Response Rate 1 is calculated.
Cooperation rate 2 calculated as the proportion of the sample successfully contacted and eligible to proceed, who completed an interview. This is consistent with the way the AAPOR Cooperation Rate 3 is calculated.

Table 2.15: SAs – Participation and response rates by visa subclass, Wave 5

Visa subclass	Eligible sample	Sample initiated	Successfully contacted & eligible to proceed	Completed interviews	Response rate 2	Cooperation rate 2
200	671	671	607	577	86.0%	95.1%
201	2	2	2	2	100.0%	100.0%
202	45	45	36	31	68.9%	86.1%
204	109	109	100	100	91.7%	100.0%
866 UMA	22	22	18	17	77.3%	94.4%
866 non-UMA	18	18	14	10	55.6%	71.4%
Total	**867**	**867**	**777**	**737**	**85.0%**	**94.9%**

Note: Formulas for calculating standardized response rates have been developed by the AAPOR: http://www.aapor.org/Education-Resources/For-Researchers/Poll-Survey-FAQ/Response-Rates-An-Overview.aspx
Response rate 2 calculated as the proportion of the sample initiated for contact who completed an interview. This is consistent with the way the AAPOR Response Rate 1 is calculated.
Cooperation rate 2 calculated as the proportion of the sample successfully contacted and eligible to proceed, who completed an interview. This is consistent with the way the AAPOR Cooperation Rate 3 is calculated.

2.5 Survey weighting

For the Wave 1 dataset, three sets of weights are provided. They were calculated to include an adjustment for non-response and take into account the probability of selection for the entire sample, as well as for PAs and SAs separately. Appendix C contains a more detailed description of the process followed to generate these weights.

For each of the subsequent waves, two types of survey weights are available: population weights and longitudinal weights. Population weights adjust BNLA estimates to population totals for PAs and SAs, as well as for a combined sample of respondents. Longitudinal weights adjust for attrition between pairs of waves of BNLA data for the same three groups. Further information on how these weights were calculated is contained in Appendices D, E, F and G.

3 Data files and variables

3.1 Data file naming conventions

The names of the available files for the BNLA general release datasets are shown in Table 3.1.

The main dataset contains the data from the PA/PR and SA/SR questionnaires. The child self-report Wave 3 dataset contains data from the child self-complete questionnaire. The accompanying BNLA Data Dictionary provides a useful summary of which questions were asked of which respondents.

Table 3.1: BNLA general release dataset file names

Wave	Main dataset name	Child self-report dataset name
1	*bnla_a20gr*	na
2	*bnla_b20gr*	na
3	*bnla_c30gr*	*bnlasr_c30gr*
4	*bnla_d40gr*	na
5	*bnla_e50gr*	na

Note: na Not applicable

3.2 Variable naming conventions

There are two unique identifier variables contained in the BNLA dataset:

- *xwaveid* is a unique 6 digit identification number (ID) assigned to every participant. This variable must be used to link unit records across waves; and
- *xmu_id* is a unique 4 digit ID assigned to every MU participating in the study—these 4 digits correspond to the first 4 digits of the unique 6 digit individual ID of the members of each MU. Anyone currently living in the same household as an existing member but who was not part of the original MU will not have an *xmu_id*.

All other variable names (besides derived and administrative variables) are made up of a number of components and adhere closely to the following convention:

> Point-in-time indicator + questionnaire module code + question number + sub-question / response option.

The point-in-time indicators are determined by when the responses are provided, that is, the time of the data collection. The point-in-time indicators, and their corresponding time of collection, can be seen in Table 3.2.

Questionnaire module codes and their corresponding topics are provided in Table 3.3.

An example of the naming convention is provided for the variable *bem15_1* (Have you found it hard getting a job for any of these reasons?) whereby:

- ***b*** is the point-in-time indicator representing a response collected at Wave 2
- ***em*** is the questionnaire module code for the Employment and income module
- ***15*** is the question number of the question in the module
- ***_1*** is question 15's Sub question / response option of "Reason hard get job – No suitable jobs suitable".

Table 3.2: BNLA Variable name point-in-time indicators and corresponding time of collection

Point-in-time indicator[(a)]	Time of data collection
z	Before permanent visa granted
a	Wave 1
b	Wave 2
c	Wave 3
d	Wave 4
e	Wave 5

(a) For easier identification, variables with the *z* point-in-time indicator also have "(Z: Pre-arrival)" in the variable label.

Table 3.3: BNLA questionnaire module codes

Module Code	Topic covered
de	Demographics
ho	Housing and neighbourhood
lp	English language proficiency
ed	Education and training
em	Employment and income
ie	Immigration experience
he	Health
ss	Self-sufficiency
cs	Community support
ls	Personal resources and life satisfaction
la	Life in Australia
cm	Child module (parent report) (Wave 3 only)
sr	Child module (child self-report) (Wave 3 only)
cg	Childcare and gender roles (Wave 5 only)

3.3 Missing data value conventions

Missing data are coded according to the following convention:

- Standard convention for "Does not apply", "Don't know" and "Prefer not to say" participant responses (explicit response options in the questionnaire):
  - –1 = Does not apply
  - –2 = Don't know
  - –3 = Prefer not to say
- Data processing
  - –4 = Not specified
  - –5 = Not asked of Principal Respondents
  - –6 = Not asked of Secondary Respondent Adults
  - –7 = Not asked of Secondary Respondent Adolescents
  - –8 = Missing data
  - –9 = Not applicable (i.e. question skipped due to answer in a preceding question)
  - –10 = Confidentialised
  - –11 = Enumerated person (i.e. a household member who did not complete a questionnaire in this wave of the study [could include study participant or non-study participant])
  - –12 = Non-Respondent/Non-Enumerated (i.e. an individual who is a study participant or was enumerated at an earlier wave of the study, but for whom no information has been collected on this occasion)

3.4 Open field responses

There were two types of open field responses: responses for the qualitative data items and the "Other (Please specify)" response where there is no obvious appropriate category. Responses to open field questions were recorded verbatim. Where necessary, these responses were translated into English by MMM translators.

Responses to the open field data items are included in the data release, but any references that could potentially identify participants have been removed from the file and replaced with a generic term in brackets. The risk of identification is expected to be low given the confidentialised status of these data. Nonetheless, data users need to be mindful at all times of their responsibility to not risk identification of respondents.

Responses entered in "Other (Please specify)" open field have been re-coded to existing categories where appropriate.

3.5 Derived items

3.5.1 Demographic characteristics and household variables

Age at interview: *<a–e>age*

Due to small numbers and to preserve anonymity, all survey respondents whose age variable had a value above 70 years were top-coded to the average age of those over 70 as follows:

- Wave 1 Respondents: 75 years; Enumerated: 78 years
- Wave 2 Respondents: 75 years; Enumerated: 79 years
- Wave 3 Respondents: 76 years; Enumerated: 78 years
- Wave 4 Respondents: 76 years; Enumerated: 79 years.
- Wave 5 Respondents: 76 years; Enumerated: 79 years

Married or has partner: *<a–e>partnered*

This variable is based on responses provided to Marital status (recorded as "Married", *<a–e>de06* = 1; or "Not married but currently has a partner", *<a–e>de07* = 1), in conjunction with responses provided by the PR for the relationship of other household members to the PR (i.e. *<a–e>de3_1*).

Where a respondent or enumerated person was identified as having a spouse, the variable *<a–e>partnered* was given a value of 1 "Yes", otherwise this variable takes a value of 0 "No".

In Wave 1, Marital/partner status was not collected for SA Adolescents; therefore, these respondents were given a value of –7 "Not asked of SA adolescent" for the variable *apartnered*. In Wave 2, SRs were not asked whether they were married (*bde06*) or whether they had a partner (*bde07*). As a result, the *bpartnered* value is based entirely on information provided by the PR in their household. With this information it is not possible to ascertain whether SRs are married/partnered to anyone other than the PR. In Wave 3, in a similar way to Wave 1, Marital/partner status was not collected for adolescent SRs; for these respondents, the value for *cpartnered* was recorded as –7 "Not asked of SR adolescent". This was not an issue from wave 4 as all SR Adolescents were of an age to be classified as SR Adults.

Country of birth

In Wave 1, PAs were asked to confirm the preloaded country of birth for each of the persons on their visa application including themselves. The responses have been coded to three distinct levels of the Standard Australian Classification of Countries (SACC) 2011[22] hierarchy for the following variables:

- *zcob_sacc* (Country of Birth)
- *zcob_minor* (Country of Birth: SACC minor groups)
- *zcob_major* (Country of Birth: SACC major groups).

[22] See ABS Standard Australian Classification of Countries (SACC), 2011 (cat. no. 1269.0)

Time between arrival in Australia and interview date: *<a–e>time_arriv*

This was calculated by subtracting the date of interview from the date of arrival held on the SDB and then grouped into categories. Neither of these two source variables has been retained in the dataset in the exact form; month and date of interview are kept in the data.

MU/Household structure: *<a-b>mu_structure/<c-d>hh_structure*

1. Couple family with children under 18 (no other family members)
2. Couple family with children under 18 and other family members
3. Couple family only (no other family members)
4. Couple family and other family members but no children under 18
5. Single parent family with children under 18 (no other family members)
6. Single parent family with children under 18 and other family members
7. Other immediate family members
8. Other extended family members only
9. Non-related persons
10. Single person

This item classifies the composition of the MU/household based on relationships to the PR. Categories were derived using the information reported by the PR which identifies the relationship of all household members to the PR. It should be noted that in deriving this item, every effort was made to provide a concise and useable classification without compromising detail. However, it was not possible to account for all permutations of the data. For example, if in an MU/household there was neither a spouse nor a biological child aged under 18 years, but there was a combination of other immediate family (i.e. parent, sibling) as well as other extended family members, then priority was given to immediate family members, and thus given a value of 7 "Other immediate family members". Other than this example, all categories are mutually exclusive. It should be noted that as the relationship is to the PR, if a member under 18 is the grandchild, niece or nephew of the PR, the family will not be classified as having children under 18.

There were no 'non-related person' relationships to the PA reported in the MU in Wave 1.

MU size: *<a–b>mu_size*

This was derived from responses provided by PA/PRs to item A1 of the questionnaire.

Total number of household members: *<a–e>tothouse*

This was derived from responses provided by PA/PRs to items A1 and A2 of the questionnaire.

Remoteness Area Name: *<a–e>raname*

Remoteness Area Name values are those that correspond to the geographical remoteness classification of the suburb of the place of residence, at the time of the BNLA interview. Firstly, the suburb of the place of residence was matched to the State Suburb Code[23] from Australian Statistical Geography Standard (ASGS) 2011. Then the State Suburb Code was

[23] See ABS Australian Statistical Geography Standard (ASGS): Volume 3 - Non ABS Structures, July 2011 (cat. no. 1270.0.55.003)

matched to the Remoteness Structure[24] of the ASGS 2011 to determine the Remoteness Area Name value.

Socio-Economic Indexes for Areas (SEIFA)

The following SEIFA 2011[25] variables are available in the data:

- *<a–e>irsad_decile* (SEIFA: Decile of Relative Socio-Economic Advantage and Disadvantage)
- *<a–e>irsd_decile* (SEIFA: Decile of Relative Socio-Economic Disadvantage)
- *<a–e>irer_decile* (SEIFA: Decile of Economic Resources)
- *<a–e>ireo_decile* (SEIFA: Decile of Education and Occupation).

SEIFA 2011 decile values are those that correspond to the suburb of the place of residence at the time of the BNLA interview.

Countries lived before Australia

PAs were asked to list the countries in which they spent time before coming to Australia. Each country nominated has been coded to three distinct levels of SACC 2011[26] hierarchy across the following variables:

- *zde10a_sacc<1–10>* (Lived before Aus - country <1–10>)
- *zde10a_minor<1–10>* (Lived before Aus - minor group <1–10>)
- *zde10a_major<1–10>* (Lived before Aus - major group <1–10>).

Number of countries lived in between country of birth and Australia: *zde10a_num*

This variable gives the value of the sum of the number of former countries of residence. It is derived using the number of countries reported in variables *zde10a_major<1–10>*. In cases where no country information was entered in these variables *zde10a_num* was coded as –4 "Not specified".

3.5.2 Language(s) spoken by respondents

Home language

In Wave 1 participants were asked to nominate the language they speak at home. A catch-up item was asked in Waves 3 and 5 for respondents who had not nominated their main language spoken at home in a previous wave. Three variables have been generated from these responses, each based upon a different level of the Australian Standard Classification of Languages (ASCL) 2011.[27] The variables are:

- *<a/c/e>lp01_lang* (Home language classification)
- *<a/c/e>lp01_narrow* (Home language narrow group classification)
- *<a/c/e>lp01_broad* (Home language broad group classification).

[24] See ABS Australian Statistical Geography Standard (ASGS): Volume 5 - Remoteness Structure, July 2011 (cat. no. 1270.0.55.005)
[25] See Census of Population and Housing: Socio-Economic Indexes for Areas (SEIFA), Australia, 2011 (cat. no. 2033.0.55.001)
[26] See ABS Standard Australian Classification of Countries (SACC), 2011 (cat. no. 1269.0)
[27] See ABS Australian Standard Classification of Languages (ASCL), 2011 (cat. no. 1267.0)

Other language(s) spoken by respondents

In Wave 3, participants were also asked to list all other languages they could speak. Again, a catch-up item was asked in Wave 5 for respondents who had not answered this item in Wave 3. Similar to the home language data item, three sets of variables have been generated from these responses based on the ASCL 2011:

- *<c/e>lp15_lang<1-3>* (Other language classification)
- *<c/e>lp15_narrow<1-3>* (Other language narrow group classification)
- *<c/e>lp15_broad<1-3>* (Other language broad group classification).

Number of languages other than English spoken by respondents

In Waves 3 and 5, a variable was derived from the responses to *<c/e>lp01_language* and *<c/e>lp15_language<1-3>*. The derived variable provides a count of the total number of languages spoken, other than English, and applies to all respondents. The variable is:

- *<c/e>lplotenum* (Number of languages other than English spoken by respondent)

3.5.3 Education and training

Main study field – before arrival (ASCED Broad field): *zed02_asced*
Main study field – current (ASCED Broad field): *<a–e>ed02_asced*

Respondents were asked whether they had undertaken study prior to coming to Australia and since being in Australia, and were then asked to record their main area of study in two open-ended items:

- *zed02_asced* (Main study area before arrival in Australia)
- *<a-e>ed02_asced* (Main study area in Australia).

Responses were analysed and coded to the minor categories of the Australian Standard Classification of Education (ASCED) 2001.[28]

3.5.4 Employment and income

Occupation before arrived (ANZSCO Sub-major groups): *zem02_<1–3>_anzsco; zem03_anzsco*

Respondents were asked to record up to three jobs they had prior to coming to Australia in three open-ended items. They were then asked to record their main occupation prior to arrival in Australia. Altogether, those four variables were:

- *zem02_1_anzsco* (Occupation before arrival in Australia – Occupation 1)
- *zem02_2_anzsco* (Occupation before arrival in Australia – Occupation 2)
- *zem02_3_anzsco* (Occupation before arrival in Australia – Occupation 3)
- *zem03_anzsco* (Main occupation before arrival in Australia).

[28] See ABS Australian Standard Classification of Education (ASCED), 2001 (cat. no. 1272.0)

Main job - Occupation (ANZSCO Sub-major group): *<a–e>em03_anzsco*

Respondents were also asked whether they had worked in the last seven days and were then asked to record their main occupation in Australia in an open-ended item: *<a–e>em03* (Main occupation in Australia).

Responses were analysed and coded to the sub-major categories of the Australian and New Zealand Standard Classification of Occupations (ANZSCO), First Edition, Revision 1.[29]

Derived weekly income (all jobs): *<a–e>em_income*

Dollar amount derived using variables *<a–e>em08* (for the only job) and *<a–e>em11* (for multiple jobs). Respondents who were unemployed were given a value of –9 "Not Applicable".

Number of financial hardship items selected: *<a–d>em26_hardship*

Derived from responses provided by PAs/PRs to items E30 *<a–f>* of the questionnaire.

3.5.5 Immigration experience

Number of refugee camps before coming to Australia: *zie13_x1*

Derived from responses provided by PAs to item F19 of the Wave 1 questionnaire.

Location of refugee camps

PAs were asked to list the countries in which they spent time in Refugee Camps before coming to Australia. Each country nominated has been coded to three distinct levels of the SACC 2011[30] hierarchy across the following variables:

- *zie14a_sacc<1–4>_x1* (Location - CAMP <1–4>)
- *zie14a_minor<1–4>_x1* (Location - CAMP <1–4> minor groups)
- *zie14a_major<1–4>_x1* (Location - CAMP <1–4> major groups).

3.5.6 Health scales

See Section 3.6.

3.5.7 Other derived variables

Religion: *acs02_religion*

The results are based upon the pre-loaded religion data obtained from the SDB in conjunction with the verbatim responses provided by some PAs to item I2 of the Wave 1 questionnaire.

In some cases members of individual MUs were classified differently to others in their MU. This occurred because only PAs were asked to respond to question I2 (What is your religion?). As such, PAs were given the opportunity to update their religion from the pre-loaded response whilst SAs were not. In addition, some PAs and SAs had different pre-loaded religions.

[29] See ABS ANZSCO - Australian and New Zealand Standard Classification of Occupations, First Edition, Revision 1 (cat. no. 1220.0)
[30] See ABS Standard Australian Classification of Countries (SACC), 2011 (cat. no. 1269.0)

Given the wide range of reasons for why MU members might have different religions, it was deemed preferable to treat all such cases as accurate, rather than change the religion categories of SAs or PAs simply to match each other.

Religion broad group classification: *acs02_broad*

Religion data were coded to the broad level of Australian Standard Classification of Religious Groups (ASGRC) 2011[31] hierarchy.

Self-esteem mean score: *als02_esteem*

Derived from responses provided to three items comprising J2 on the Wave 1 questionnaire (variables *als02<a >)*. The score was derived by taking an average of responses to the three variables. Lower values of the score refer to better self-esteem. The score was derived for all respondents who have provided a valid response to at least one item in the scale.

Self-efficacy mean score: *<a/c>ls03_efficacy*

Derived from responses provided to the three items comprising J3 on the Wave 1, 3 and 5 questionnaires (variables *<a/c/e>ls03<a–c>)*. The score was derived by taking an average of responses to the three variables. Lower values of the score refer to greater self-efficacy. The score was derived for all respondents who have provided a valid response to at least one item in the scale.

3.6 Health scales

The questionnaires include health scales that can be compared back to population norms. The following scales are used:

- Kessler 6 (K6) psychological distress[32]
- 8-item posttraumatic stress disorder (PTSD-8)[33]
- SDQ for parents or teachers of 4–17 year olds, UK English version[34]
- SDQ self-rated for 11–17 year olds, Australia English version.[35]

The item variables associated with each of the health scales are shown in Table 3.4.

Table 3.4: BNLA item variables for K6, PTSD-8 and SDQ scales

Scale	BNLA Variables
K6	*<a–e>he08<a–f>*
PTSD-8	*<a–e>he10<a–h>* *csr<37-45>*
SDQ (for parents or teachers of 4–17 year olds)	*ccm01<a–y>_1* *ccm01<a–y>_2*
SDQ (self-rated for 11–17 year olds)	*csr<01–25>*

[31] See ABS Australian Standard Classification of Religious Groups, 2011 (cat. no. 1266.0)
[32] RC Kessler, PR Barker, LJ Colpe, JF Epstein, JC Gfroerer, E Hiripi, MJ Howes, SL Normand, RW Manderscheid, EE Walters & AM Zaslavsky, 'Screening for Serious Mental Illness in the General Population', *Archives of General Psychiatry*, vol. 60, no. 2, 2003, pp. 184-189.
[33] M Hansen, TE Anderson, C Armour, A Elklit, S Palic & T Mackrill, 'PTSD-8: A Short PTSD Inventory', *Clinical Practice & Epidemiology in Mental Health*, vol. 6, 2010, pp. 101-108.
[34] youthinmind, *One-sided SDQ for parents or teachers of 4-17 year olds,* retrieved 18 April 2017 from http://www.sdqinfo.org/py/sdqinfo/b3.py?language=Englishqz(UK)
[35] youthinmind, *One-sided self-rated SDQ for 11-17 year olds,* retrieved 1 March 2017 from http://www.sdqinfo.org/py/sdqinfo/b3.py?language=Englishqz(Austral)

3.6.1 K6

The K6 is a measure of psychological distress. K6 score groups provide an indication of whether a severe mental illness is likely to be present or not. The scores should not be interpreted as a diagnosis of (or lack of) a mental illness.[36]

Values for the K6 Total Score—variable *<a–e>he08_k6score*—are calculated by summing individual scores across all items in variables *<a–e>he08<a–f>* in accordance with the standard scoring method:

1 = None of the time
2 = A little of the time
3 = Some of the time
4 = Most of the time
5 = All of the time

Values for the K6 Score Group—variable *<a–e>he08_k6group*—are then derived from the *<a–e>he08_k6score* values, in accordance with the standard K6 Score Group method:

1 = Total score 6–18 (No probable serious mental illness)
2 = Total score 19–30 (Probable serious mental illness)

3.6.2 PTSD-8

Hansen et al. suggest one way to score the PTSD-8 scale.[37] The criteria for indicating the presence of posttraumatic stress disorder (PTSD) are met if at least one symptom from each of the three PTSD-8 subscales has an item score of 3 or 4 (i.e. respondents answered either 3 "Sometimes" or 4 "Most of the time" for at least one item in each subscale). The three subscales are:

- Intrusion
- Avoidance
- Hypervigilance.

The item variables from which each subscale is derived are listed in Table 3.5. The derived PTSD-8 variables are:

- *<a–e>he10_ptsd8* for the main dataset
- *csrptsd8* for the child self-report dataset.

Due to the non-clinical nature of the assessment, data users are advised to treat these PTSD-8 results as indicative only. The PTSD-8 was originally designed to identify symptoms experienced between the time of trauma and the time of interview. In BNLA the questions were asked in relation to the previous week. Additionally, all respondents were asked the questions regardless of whether they had experienced traumatic events. Due to the difference in measurement methodology, the measure in BNLA more accurately reflects whether some

[36] Further information on scoring of the Kessler-6 can be found at http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/4817.0.55.001Chapter92007-08
[37] M Hansen, TE Anderson, C Armour, A Elklit, S Palic & T Mackrill, 'PTSD-8: A Short PTSD Inventory', *Clinical Practice & Epidemiology in Mental Health*, vol. 6, 2010, pp. 101-108.

has experienced symptoms of PTSD in the previous week rather than whether they meet the criteria for PTSD.

Table 3.5: BNLA PTSD-8 subscale item variables for main dataset and child self-report dataset

PTSD-8 subscale	Main dataset item variables	Child self-report dataset item variables
Intrusion	*<a–e>he10a* *<a–e>he10b* *<a–e>he10c* *<a–e>he10h*	*csr38* *csr39* *csr40* *csr45*
Avoidance	*<a–e>he10f* *<a–e>he10g*	*csr43* *csr44*
Hypervigilance	*<a–e>he10d* *<a–e>he10e*	*csr41* *csr42*

Treatment of missing values in PTSD-8 data items in the child self-report dataset

Due to the pen and paper mode of data collection for the child self-report instrument, there was a small number of missing responses to one or more items in the PTSD-8 scale. In such cases, –4 "Not specified" has been recorded in the data. To derive the PTSD-8 indicator in the presence of missing values, the following protocol was adopted:

- For the Intrusion subscale, the symptom cluster criterion was met if at least one item had a score of 3 or 4 regardless of the number of missing responses. If the symptom cluster criterion was not met and one of the four items' responses was missing, the symptom cluster was scored as not meeting the criterion. If the symptom cluster criterion was not met and two or more of the four items' responses were missing, the value of the symptom cluster was set to missing.
- For the Avoidance and Hypervigilance subscales, the symptom cluster criterion was met if at least one item had a score of 3 or 4 regardless of the number of missing responses. If the symptom cluster criterion was not met and if one or both of its two items' responses were missing, the value of the symptom cluster was set to missing.
- For the overall PTSD-8 measure:
    - if all of the subscales' criteria were met, the PTSD-8 value was set to 'May have PTSD'
    - if all three subscales had a value of missing, the PTSD-8 value was set to missing ("Not specified")
    - if any of the subscales' criteria was not met, the PTSD-8 value was set to 'Unlikely to have PTSD'
    - if one or two subscales' criteria were met and the remaining subscale(s) had a value of missing, then the PTSD-8 value was set to missing ("Not specified").

3.6.3 SDQ[38]

The Strengths and Difficulties Questionnaire (SDQ) is a brief behavioural screening questionnaire about children and young people aged 3 to 16 years. It consists of 25 items with three response options: "Not True"; "Somewhat True"; and "Certainly True". The items comprise the following five scales of five items each:

- Prosocial
- Emotional Symptoms
- Conduct Problems
- Hyperactivity
- Peer Problems.

For each of the five scales, the score can range from 0 to 10 if all items were completed. These scores can be scaled up pro-rata if at least three items were completed.

The SDQ Total Difficulties Score is generated by summing scores from all the scales except the Prosocial scale. The resultant score has a range of 0 to 40 and is counted as missing if one of the four component scores is missing.[39] Higher values of the SDQ Total Difficulties Score indicate a higher risk of social-emotional problems. For the SDQ Prosocial Score, higher values indicate a higher likelihood of undertaking "voluntary actions that are intended to help or benefit another individual or group of individuals".[40]

Wave 3 BNLA datasets contain a derived variable for each of the five SDQ scales.

Table 3.6 presents the derived variables for the SDQ scales and their associated item variables, separately for the main dataset and the child self-report dataset.

[38] © Robert Goodman
[39] youthinmind, *English (Austral) Scoring instructions for SDQs for 4-17 year olds, completed by parents, teachers or self-report [and for those aged 18+],* 11 May 2015, retrieved 24 February 2017 from http://www.sdqinfo.org/py/sdqinfo/b3.py?language=Englishqz(Austral)
[40] N Eisenberg & PH Mussen, *The Roots of Prosocial Behavior in Children*, Cambridge University Press, Cambridge, 1989, p. 3.

Table 3.6: SDQ scale derived variables and item variables for BNLA

SDQ scale	Main dataset derived variable	Main dataset item variables	Child self-report dataset derived variable	Child self-report dataset item variables
Prosocial	*ccm01pros_<1–2>*	*ccm01a_<1–2>* *ccm01d_<1–2>* *ccm01i_<1–2>* *ccm01q_<1–2>* *ccm01t_<1–2>*	*csrpros*	*csr01* *csr04* *csr09* *csr17* *csr20*
Emotional Symptoms	*ccm01emot_<1–2>*	*ccm01c_<1–2>* *ccm01h_<1–2>* *ccm01m_<1–2>* *ccm01p_<1–2>* *ccm01x_<1–2>*	*csremot*	*csr03* *csr08* *csr13* *csr16* *csr24*
Conduct Problems	*ccm01cond_<1–2>*	*ccm01e_<1–2>* *ccm01g_<1–2>** *ccm01l_<1–2>* *ccm01r_<1–2>* *ccm01v_<1–2>*	*csrcond*	*csr05* *csr07** *csr12* *csr18* *csr22*
Hyperactivity	*ccm01hype_<1–2>*	*ccm01b_<1–2>* *ccm01j_<1–2>* *ccm01o_<1–2>* *ccm01u_<1–2>** *ccm01y_<1–2>**	*csrhype*	*csr02* *csr10* *csr15* *csr21** *csr25**
Peer Problems	*ccm01peer_<1–2>*	*ccm01f_<1–2>* *ccm01k_<1–2>** *ccm01n_<1–2>** *ccm01s_<1–2>* *ccm01w_<1–2>*	*csrpeer*	*csr06* *csr11** *csr14** *csr19* *csr23*

Note: * indicates items reversed for scoring.

3.7 Data amendments for confidentiality reasons

Some items that were asked in the survey have not been included in the dataset as these variables could potentially identify the respondent:

- Respondent's date of birth (question A1e in the MU/household grid – see PA/PR survey); however, a derived variable, *<a–e>age*, provides the respondent's age.
- Respondent's ethnicity (question I1).
- Year of arrival in former countries of residence (question A10b – see PA survey). Due to the diverse migration pathways of respondents, this information is potentiality identifying when used in conjunction with other data in question A10 (i.e. country name and time spent in each country). However, we have retained these other variables in the dataset: *zde10a** (Country name); and *zde10c** (Number of years and/or months spent in each country). This is sufficient to conduct meaningful analyses of these data.

The following items have also been amended to assist with retaining respondents' anonymity:

- Age (*<a–e>age*) – Due to small numbers and to preserve anonymity, all responses where the age variable had a value above 70 years were coded to the average age of participants. The averages were calculated separately for responding and enumerated persons (see Section 3.5.1).
- Variables associated with country of birth (*zcob_sacc* and *zcob_minor*), countries lived in (*zde10a_sacc<1–10>* and *zde10a_minor<1–10>*) and locations of refugee camps (*zie14a_sacc<1–4>_x1* and *zie14a_minor<1–4>_x1*), have also been subject to confidentialisation. In these variables, countries for which there were fewer than 10 households with a member who nominated it were reclassified as –10 "Confidentialised".
- Variables associated with home languages of respondents, other languages spoken by respondents, or language of interview have also been subject to confidentialisation. In these variables, languages and narrow groups of languages which had fewer than 10 respondents were reclassified as –10 "Confidentialised".
- Religions which were nominated by fewer than 10 households in *acs02_religion* have also been coded as –10 "Confidentialised".

See also Section 3.4 for information about treatment of open field responses.

3.8 Disability status and NDIS participation

In Wave 5, two new items were added to the PaRMS administrative household database to capture information for each household member regarding:

- Disability status (*ehhdis*)*: "*Do you/does this person have a disability, injury or health condition that has lasted or is likely to last 12 months or more?"; and
- Participation in the National Disability Insurance Scheme (*ehhndis)*: "Are you/is this person a participant on the National Disability Insurance Scheme?"

In previous waves, respondents were asked about their own disability status within the survey *(<a-d>he09*) using the same wording used in *ehhdis.* For longitudinal analysis, *ehhdis* can be analysed with *<a-d>he09* for data users who are interested in looking at respondent's disability status over time. Note that in Wave 5 the PR was asked the questions in relation to all members of the household and the original MU.

3.9 Other notes relevant to the data

3.8.1 Changes since Release 3.0

Starting from Release 4.0, some household member information is presented in both a household-level ('wide') format and a person-level ('long') format.

The datasets are presented in a person-level format, with one observation for each respondent and each non-responding person. Non-respondents can be identified in the dataset as an "Enumerated person" or "Non-responding/Non-enumerated person" in the variable *<a–e>resp*.

The following person-level variables have been derived for each person in the dataset:

- *<a–e>prid* – xwaveid of the Principal Respondent in the Migrating Unit/household
- *<a–e>hhpno* – member number in the Migrating Unit/household
- *<a–e>hhrel* – relationship to the Principal Respondent in the Migrating Unit/household
- *<a–e>hhlive* – whether the Respondent lives in the same household as the Principal Applicant on the original visa application.

In addition, variables listing other members of the Migrating Unit/household are available for each person. These are derived from PR responses to item A1 of the PR instrument, and are the same for all members of the Migrating Unit/household. These 'wide'-format variables are:

- *<a-e>de01a_X* – xwaveid of MU/household member X
- *<a-e>de0b_X* – gender of MU/household member X
- *<a-e>de01c_X* – age of MU/household member X
- *<a-e>de01d_X* – applicant type (pre-arrival) of MU/household member X
- *<a-e>de01e_X* – country of birth of MU/household member X
- *<a-e>de01f_X* – relationship to PR of MU/household member X
- *<a-e>de01g_X* – whether member X has a disability, long-term injury or health condition
- *<a-e>de01h_X* – whether member X is a participant of NDIS.

Another change since Release 3.0 was the addition of code 3 "Deceased" to the response categories for *<c-e>hhlive*.

3.8.2 Self-reported data

Data users should note that all data are self-reported. This may explain some slightly anomalous data such as the small number of participants who reported far higher levels of government income support payments than expected (see variable *<a-e>em19*).

3.8.3 Missing data

For a number of numeric variables the survey programming did not force a response to be entered on the screen (i.e. the respondent could move on to the next question without entering a response). In such cases, –4 "Not specified" has been recorded. All other questions besides verbatim responses were formatted to ensure that a response was provided.

For a small number of variables in Wave 1, data was not collected where it would otherwise have been expected based on the sequencing of the survey. This arose due to a technical issue with the administration of the survey which allowed respondents to skip through survey screens. The number of cases with missing data for this reason ranged from one to five cases across these variables. The variables affected by this issue are listed in Table 3.7.

Due to an error in the sequencing of respondents in the Wave 1 Employment section, only a small proportion of eligible respondents were asked the following two questions:

- E22 Even though you are not currently working, would you like to get a job?
- E23 What are the reasons you do not want to get a job? (mark all that apply)

As a result, responses to these two questions have been removed from the dataset.

In Wave 4, a small number of respondents had missing data in questions related to partner or children because of incorrect or not up-to-date information in the MU grid. In these cases, the information in the MU grid was updated after the survey had been completed. The variables affected and the number of missing cases are shown in Table 3.8.

3.8.4 Relationship to PA/PR data item

Information concerning household members' relationship to the PA/PR was obtained from the PA/PR in all waves. In Wave 1 relationship information was obtained from the SDB and pre-populated into the questionnaire. Prior to conducting the interview, all such information was updated by the PA (i.e. name, age, gender, country of birth and relationship to PA for all MU members). In a small number of cases (n = 49 MUs), the interviewer did not update all information in the MU grid. In these cases, the MU grid data contains the original administrative information supplied in the SDB. Efforts were made to re-contact relevant MUs to update this information with the correct relationship classification; however, this was not possible in all cases. It is for this reason that some relationship data are coded as "Dependent", without further classification.

Table 3.7: Variables skipped in error in Wave 1 due to technical issue

Variables			
ade06	*aho08<a–g>*	*aem13b1*	*ass01a*
ade07	*alp09*	*aem13b2*	*ass01<b–i>*
ade08	*alp10_<1–3>*	*aem13b3*	*acs03<a–c>*
ade11	*alp13_<1–5>*	*ahe11_7*	*acs08d*

Table 3.8: Missing data in partner and children questions in Wave 4

Variable	Number of missing cases
dde07	1
dhe11_8	24
dss01c	24
dcs09a-dcs09g	24
dls13a-dls13c	1

3.8.5 Interview dates

Data users should be aware that although technical procedures[41] were in place during Wave 1 fieldwork to ensure that respondents in the offshore group were not interviewed less than three months after their arrival in Australia, there are 39 offshore arrivals recorded as having arrived less than three months earlier. Investigations into this reveal a number of distinct explanations including:

- Some interviewers updated the regional settings on their tablet computers, which would have reset the time zone and date format resulting in an incorrect interview date being recorded.
- Some interviewers may have accessed the survey directly through Survey Centre[42] rather than Participant and Respondent Management System (PaRMS), which would have overridden the technical checks in place in PaRMS.

However, it is clear that if any individuals were interviewed before they had been in Australia for three months it would have been no more than a few days before their eligibility date.

3.8.6 Length of time in Australia

Data users should also be aware that respondents in the onshore group (e.g. visa subclass 866) have been in Australia for longer on average than their offshore counterparts. As discussed in Section 2, to be eligible for the study, offshore visa holders had to have arrived in Australia holding a permanent visa three to six months prior to their Wave 1 interview, whereas onshore visa holders had to have received their permanent protection visa in the same three to six month period prior to Wave 1 to be eligible for the study. It is important to be mindful of this when making data comparisons between the onshore and offshore groups.

The length of time between arrival in Australia and Wave 1 interview is detailed in Table 3.9. A minimum of 9 months was required to elapse between each main wave interview after Wave 1 (see also Section 2.2.6).

3.8.7 PR Report responses

For a number of variables, responses provided by PRs have been copied across to the corresponding fields of SRs in their respective MUs. This occurred for those questions which were asked of PRs but not of SRs and for which the responses provided by PRs were deemed to be relevant to all the SRs in their MU (e.g. housing in Australia, some migration pathways and financial hardship). Variables for which this has been undertaken can be identified in two ways:

- the variable's label in all such cases includes the text "(PR Report)"
- the Data Dictionary identifies such variables in the *w<1–5>PRreport* field as "Yes".

[41] Surveys are accessed primarily through the Participant and Respondent Management System (PaRMS). Logic was built into the PaRMS system to ensure that records were not activated or released to interviewers until three months after arrival for offshore respondents, and three months after visa grant date for onshore respondents.
[42] Survey Centre is the software into which responses to survey questions were recorded directly into the tablet computer.

Table 3.9: Number of days between date of arrival in Australia and Wave 1 interview, by migration pathway – number of respondents

Number of days between arrival in Australia and Wave 1 interview	Migration Pathway		
	Onshore	Offshore	Total
Less than 3 months	0	39	39
3 to less than 6 months	0	1,766	1,766
6 to less than 12 months	12	214	226
1 to less than 2 years	283	0	283
2 to less than 3 years	40	0	40
3 or more years	45	0	45
Total	380	2,019	2,399

Note: Data are unweighted.
95% of offshore respondents categorised as having been in Australia for 6 to 12 months had only been in Australia for approximately 6 months and 1 week.

3.10 Data access and further information

The General Release BNLA datasets are available to download, free of charge, through the Dataverse platform hosted by the Australian Data Archive (ADA). There are strict security and confidentiality protocols surrounding use of the data. Prospective users are required to complete a Data Access Request. Before applying for access to the data, researchers should make themselves familiar with the NCLD Access and Use Protocols. These are available at the ADA website at https://dataverse.ada.edu.au/dataverse/ncld.

More information about the study can be found at the BNLA website: http://www3.aifs.gov.au/bnla/

Enquires can also be sent via email to either: bnla@aifs.gov.au or NCLDResearch@dss.gov.au

Appendix A: Sample characteristics

The distribution of country of birth for the BNLA eligible sample is described in Table A.1.

Table A.1: BNLA eligible sample – Country of birth

Country of birth	Proportion of sample (%)
Afghanistan	23.8
Bhutan	2.3
Congo, Democratic Republic of	1.4
Egypt	3.4
Eritrea	1.5
Ethiopia	2.0
India	0.6
Iran	10.0
Iraq	25.8
Libya	0.8
Myanmar (Burma)	12.0
Nepal	0.1
Pakistan	4.6
Sri Lanka	2.5
Sudan	0.6
Syria	1.3
Other - Confidentialised	7.3
Total	**100.0**

Source: DIBP SDB

Appendix B: Survey completion overview

A sample size of 1,500 PAs was the target for Wave 1 of the study, with no constraints on the number of SAs per MU. The initial sample targets were for 70% offshore and 30% onshore, with all eligible MUs to be contacted (i.e. a census approach). A quota approach was to be used to recruit the onshore cohort (given the considerably larger numbers expected to be available than needed). These parameters were subsequently revised due to the implementation of offshore processing in the lead up to the main wave. This policy change resulted in a sudden decline in the number of persons being granted onshore protection visas. Targets were adjusted to accommodate the lower number of onshore MUs anticipated to be available, with a census approach instead of a quota approach used to recruit the onshore sample. As seen in Table B.1, the 866 UMA quota was reduced and the deficit was made up in the 200 and 204 visa subgroups. The revised sample targets were 77% offshore and 23% onshore MUs (i.e. at the PA level).

For Wave 2 onwards, the quota values were calculated using the number of achieved interviews in Wave 1 as the base value and an assumed 10% attrition rate for each subsequent wave. For example, for the quota value for the 200 visa class in Wave 2, based on 942 achieved interviews in Wave 1 and an assumed attrition rate of 10%, meant that the quota for Wave 2 was 848 (i.e. 847.8 rounded to the nearest whole number). Then, based on the unrounded quota value of 847.8 in Wave 2 and an assumed attrition rate of 10%, the quota for the 200 visa class for Wave 3 was 736 (i.e. 736.02 rounded to the nearest whole number).

Prior to the commencement of Wave 4, new sample targets were established. The new sample targets aimed to achieve the number of interviews completed in the previous wave. Based on Wave 3 achieved interviews, the Wave 4 sample targets were: 1,188 PAs and 706 SAs. Adjusted sample targets were also established for Wave 5. The decision was made to define the target as the number of interviews completed in the previous face-to-face wave (Wave 3). Based on the Wave 3 achieved interviews, the Wave 5 sample targets were 1,188 PAs and 706 SAs.

Table B.1: Wave 1 MU targets by visa subclass

Visa subclass	Original quota		Revised quota	
	Number of PAs	Proportion of sample (%)	Number of PAs	Proportion of sample (%)
Offshore	**1,050**	**70.0**	**1,160**	**77.0**
200	850	56.7	940	62.7
201	8	<1	8	<1
202	40	2.7	40	2.7
203	2	<1	2	<1
204	150	10.0	170	11.3
Offshore	**450**	**30.0**	**340**	**23.0**
866 UMA	360	24.0	250	16.7
866 non-UMA	90	6.0	90	6.0
Total	**1,500**	**100.0**	**1,500**	**100.0**

B.1 Wave 1 survey completion

The following three tables describe the gender breakdown of participants in each visa class subgroup for Wave 1. Each table focuses on one of the three respondent types: PAs (Table B.2); SA Adults (Table B.3); or SA Adolescents (Table B.4).

The majority of PAs (70%; *n* = 1061) in Wave 1 were male (Table B.2). In contrast, most SA Adults were female (75%; *n* = 567; Table B.3). SA Adolescents were more evenly divided between males and females (57% and 43% respectively; Table B.4).

As can be seen in Table B.2, the majority of PAs were male in visa class 200 and in both onshore visa classes 866 UMA and 866 non-UMA. Unsurprisingly, visa class 204 (Woman at Risk) is overwhelmingly made up of women as this category provides asylum to women who do not have the protection of a male relative and are in danger of victimisation, harassment or serious abuse because of their gender.[43]

Although the majority of people who are issued 204 visas are female, a small number of males also enter Australia under this visa class as children or other family members of a primary applicant who is female. This occurs when a subclass 204 visa holder proposes a son or other family member who happens to be male. They then receive the same visa subclass as the 204 proposer. This is the case for the *n* = 38 male survey participants in this visa class.

Table B.2: PAs – Survey completion, by visa subclass and gender, Wave 1

PA interviews achieved	Offshore					Onshore		Total
	200	201	202	203	204	866 UMA	866 non-UMA	
Gender								
Male	745	5	16	0	3	229	63	1,061
Female	197	3	26	0	180	5	37	448
Total	942	8	42	0	183	234	100	1,509
Quota	940	8	40	2	170	250	90	1,500
Proportion of quota	100%	100%	105%	0%	108%	94%	111%	101%

[43] Department of Immigration and Border Protection (2014) Woman at risk visa (subclass 204), DIBP website, http://www.immi.gov.au/Visas/Pages/204.aspx Accessed 29 April 2014.

Table B.3: SA Adults – Survey completion, by visa subclass and gender, Wave 1

SA Adult interviews achieved	Offshore					Onshore		Total
	200	201	202	203	204	866 UMA	866 non-UMA	
Gender								
Male	148	2	9	0	23	3	3	188
Female	453	0	22	0	55	21	16	567
Total	601	2	31	0	78	24	19	755

Table B.4: SA Adolescents – Survey completion, by visa subclass and gender, Wave 1

SA Adolescent interviews achieved	Offshore					Onshore		Total
	200	201	202	203	204	866 UMA	866 non-UMA	
Gender								
Male	40	0	3	0	12	2	1	58
Female	45	0	13	0	19	0	0	77
Total	85	0	16	0	31	2	1	135

B.3 Wave 2 survey completion

In Wave 2, interviews were achieved with 1,268 PAs (84% of Wave 1), 626 SA Adults (83% of Wave 1) and 115 SA Adolescents (85% of Wave 1). Whilst the 90% quota target for PAs was not achieved in Wave 2, all visa subclasses (excluding visa 204) were within 90% to 97% of the target. As previously outlined, the largest reason for non-response was due to inability to make contact.

The following three tables describe the gender breakdown of participants in each visa class subgroup for Wave 2. Each table focuses on one of the three respondent types: PAs (Table B.5); SA Adults (Table B.6); or SA Adolescents (Table B.7).

Table B.5: PAs – Survey completion, by visa subclass and gender, Wave 2

PA interviews achieved	Offshore				Onshore		Total
	200	201	202	204	866 UMA	866 non-UMA	
Gender							
Male	636	4	15	2	192	52	901
Female	171	3	20	139	4	30	367
Total	807	7	35	141	196	82	1,268
Quota	848	7	38	165	211	90	1,358
Proportion of quota	95%	97%	93%	86%	93%	91%	93%

Table B.6: SA Adults – Survey completion, by visa subclass and gender, Wave 2

SA Adult interviews achieved	Offshore				Onshore		Total
	200	201	202	204	866 UMA	866 non-UMA	
Gender							
Male	131	2	5	20	3	2	163
Female	374	0	15	47	16	11	463
Total	505	2	20	67	19	13	626

Table B.7: SA Adolescents – Survey completion, by visa subclass and gender, Wave 2

SA Adolescent interviews achieved	Offshore			Onshore	Total
	200	202	204	866 UMA	
Gender					
Male	36	3	7	2	48
Female	39	10	18	0	67
Total	75	13	25	2	115

B.3 Wave 3 survey completion

In Wave 3, interviews were achieved with 1,155 PAs (77% of Wave 1), 624 SA Adults (83% of Wave 1) and 115 SA Adolescents (85% of Wave 1). Overall, 95% of the quota target for PAs was achieved in Wave 3, with participants holding a 200, 201 and 204 visa subclass achieving the highest percentage of the quota (between 99% and 107%).

The following three tables describe the gender breakdown of participants in each visa class subgroup for Wave 3. Each table focuses on one of the three respondent types: PAs (Table B.8); SA Adults (Table B.9); or SA Adolescents (Table B.10).

Table B.8: PAs – Survey completion, by visa subclass and gender, Wave 3

PA interviews achieved	Offshore				Onshore		Total
	200	201	202	204	866 UMA	866 non-UMA	
Gender							
Male	591	4	14	2	157	38	806
Female	161	2	14	145	3	24	349
Total	752	6	28	147	160	62	1,155
Quota	763	6	34	148	190	81	1,222
Proportion of quota	99%	100%	82%	107%	84%	77%	95%

Table B.9: SA Adults – Survey completion, by visa subclass and gender, Wave 3

SA Adult interviews achieved	Offshore				Onshore		Total
	200	201	202	204	866 UMA	866 non-UMA	
Gender							
Male	123	1	5	21	1	2	153
Female	389	0	15	46	12	9	471
Total	512	1	20	67	13	11	624

Table B.10: SA Adolescents – Survey completion, by visa subclass and gender, Wave 3

SA Adolescent interviews achieved	Offshore			Onshore	Total
	200	202	204	866 UMA	
Gender					
Male	33	3	10	1	47
Female	41	10	17	0	68
Total	74	13	27	1	115

B.4 Wave 4 survey completion

In Wave 4, interviews were achieved with 1,186 PAs (79% of Wave 1) and 743 SAs (83% of Wave 1, including adolescents). The following two tables show survey completion by gender and visa subclass for Wave 4 (Table B.11 for PAs and Table B.12 for SAs).

Table B.11: PAs – Survey completion, by visa subclass and gender, Wave 4

PA interviews achieved	Offshore				Onshore		Total
	200	201	202	204	866 UMA	866 non-UMA	
Gender							
Male	162	3	15	151	3	24	358
Female	622	4	14	3	143	42	828
Total	**784**	**7**	**29**	**154**	**146**	**66**	**1,186**

Table B.12: SA Adults – Survey completion, by visa subclass and gender, Wave 4

SA interviews achieved	Offshore				Onshore		Total
	200	201	202	204	866 UMA	866 non-UMA	
Gender							
Male	430	0	24	63	12	10	539
Female	157	2	7	31	5	2	204
Total	**587**	**2**	**31**	**94**	**17**	**12**	**743**

B.5 Wave 5 survey completion

In Wave 5, a total of 1,881 interviews were achieved, which included 1,144 PAs (76% of Wave 1) and 737 SAs (83% of Wave 1). Inability to contact respondents was the largest reason for non-response, with only around 5% of respondents who were successfully contacted in Wave 5 refusing an interview. The following two tables show survey completion by gender and visa subclass for Wave 5 (Table B.13 for PAs and Table B.14 for SAs).

Table B.13: PAs – Survey completion, by visa subclass and gender, Wave 5

PA interviews achieved	Offshore				Onshore		Total
	200	201	202	204	866 UMA	866 non-UMA	
Gender							
Male	598	4	15	2	148	30	797
Female	155	3	17	147	2	23	347
Total	**753**	**7**	**32**	**149**	**150**	**53**	**1,144**

Table B.14: SAs – Survey completion, by visa subclass and gender, Wave 5

SA interviews achieved	Offshore				Onshore		Total
	200	201	202	204	866 UMA	866 non-UMA	
Gender							
Male	154	2	8	31	2	2	199
Female	423		23	69	15	8	538
Total	**577**	**2**	**31**	**100**	**17**	**10**	**737**

Appendix C: Wave 1 survey weighting

An analysis of non-response was undertaken where the characteristics of the Wave 1 participating individuals were compared with all eligible participants on the DIBP sample frame (n = 7,362). A range of factors were considered in this analysis including gender, visa subclass, age, location of residence, family size, marital status and country of birth.

Multivariate logistic models were then fitted to identify the factors associated with survey completion, with two separate models being fitted. The first model explored the factors predicting participation of PAs, the second model examined the characteristics related to participation by SAs.

Significant variables predicting PA participation included:

- visa subclass;
- location of residence;
- family size; and
- country of birth.

For SAs, survey completion was significantly associated with:

- gender;
- visa subclass;
- location of residence;
- age;
- country of birth; and
- married females in a household with a male PA.

Compared to the eligible PA sample, the survey had lower representation of PAs living in capital cities. Other factors associated with a lower probability of participation included smaller sized families, participants with a non-UMA visa subclass[44] and those born in Burma, the rest of Asia or Africa.

Analysis at the SA level revealed that being older, a male SA, living in a capital city, having a non-UMA visa subclass, being born in Burma, the rest of Asia or Africa and not being a married female in a household with a male PA, were all significantly associated with lower levels of survey participation.

The analysis above identified that different factors explained survey completion between PAs and SAs. To further investigate the appropriateness of calculating a survey weight across both PA and SA participants a variance components model was also estimated. A model with and without covariates was fitted. The results from the unadjusted model showed that only around 11% of the variance in survey completion was explained at the individual level (i.e. SA level). A model with the covariates found significant in the PA analysis described earlier was also fitted.[45] These variable inclusions added little further explanatory power when compared to the unadjusted model. Given the low level of variance explained at the SA level, these results

[44] A quota was in place for participants with a non-UMA visa subclass.
[45] Family size was excluded from the variance components model as this variable was not statistically significant when this model was estimated. Age was also included as an explanatory variable in this model.

suggested it was appropriate to also calculate a single survey weight covering both PAs and SAs (variable: *awgt*), in addition to a PA level (*apawgt*) and SA level weight (*asawgt*).

To adjust for the differential pattern in non-response described above, survey weights were calculated using the *ipfweight* algorithm (also known as raking) in Stata authored by Michael Bergmann.[46] The *ipfweight* procedure is based on work first proposed by Deming and Stephan (1940)[47] and adjusts survey sample weights to achieve population totals. In this case, the population was of humanitarian migrants being granted a permanent visa between May and December 2013.[48] Three survey weights were calculated using this procedure and are available for use in the BNLA Wave 1 dataset. The survey weights and their variable names are described further below.

1. ***apawgt*** – PA survey weight. The following variables were used in the weighting procedure for this variable: visa subclass, capital city, family size and country of birth. The generated survey weight ranged in value from 0.40 to 4.11. After examining the distribution and factors associated with higher survey weights, the calculated survey weights were capped at a maximum value of 2.5 (27 out of 1,509 PAs in Wave 1 had an initial survey weight greater than 2.5). It is recommended that this survey weight be used for analysis involving only questions asked of PAs.

2. ***asawgt*** – SA survey weight. The gender, visa subclass, capital city, age, country of birth and married female variables were used in the weighting procedure for this variable. Survey weights ranged in value from 0.47 to 6.35. As with the PA level weight, the SA level survey weight was constrained to take on a maximum value of 2.5 (involving 13 out of 890 cases). It is recommended that this survey weight be used for analysis involving only questions asked of SAs.

3. ***awgt*** – all participant survey weight. Gender, visa subclass, capital city, age and country of birth information were used in the weighting for this variable. This survey weight ranged from 0.37 to 4.28 and was also capped at a maximum value of 2.5 (involving 37 out of 2,399 cases). It is recommended that this survey weight be used when analysing survey questions asked of both PA and SA respondents.

   While not specifically designed for this purpose, the all participant survey weight (*awgt*) is the most appropriate survey weight to use when analysing variables for which responses provided by PAs have been copied across to the corresponding SAs in their MU.

[46] See http://fmwww.bc.edu/repec/bocode/i/ipfweight.html for further detail. Accessed on 29/4/2014.
[47] Deming, W. Edwards/Frederick F. Stephan. 1940. On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals Are Known, in: *The Annals of Mathematical Statistics* 11 (4): 427-444.
[48] Offshore visa holders had to have **arrived** in Australia holding a permanent visa, while onshore visa holders had to have **received** their protection visa in this time period.

Appendix D: Wave 2 survey weighting

This section describes the process used to calculate survey weights for the Wave 2 sample of BNLA. Two types of survey weights were calculated and are available on the Wave 2 dataset:

1. A population weight, which is a variable that adjusts BNLA estimates to population totals (in this case, the population of humanitarian migrants who have been granted a permanent visa or arrived in Australia between May and December 2013).[49]
2. A longitudinal weight, which is a variable that adjusts for attrition between the first and second wave of BNLA data.

D.1 Calculation of Wave 2 population weights

The same process was used to calculate Wave 2 population weights as was followed for Wave 1 survey weighting. Appendix C describes this methodology more fully, but in brief, an analysis of non-response was undertaken to compare the characteristics of the Wave 2 participating sample with all eligible PAs and SAs participants on the DIBP sample frame.[50,51] This was done using multivariate logistic models to identify the factors associated with Wave 2 survey completion. As in Wave 1, two separate statistical models were fitted – the first explored the factors predicting participation of PA participants and the second model identified factors associated with SA participation. A third model was also fitted which examined the characteristics associated with survey completion of all participants.

D.1.1 PA participation in Wave 2

The multivariate logistic model results predicting Wave 2 PA participation are described in Table D.1, which has a breakdown of potential BNLA participants by characteristics drawn from the SDB. Based on the model results, statistically significant variables predicting PA participation included:

- visa subclass,
- region,
- family size, and
- country of birth.

[49] Offshore visa holders had to have **arrived** in Australia holding a permanent visa, while onshore visa holders had to have **received** their protection visa in this time period to be eligible for participation in the BNLA.
[50] *n* =7,362 individuals in 4,035 MUs were eligible for selection. For more information about the sampling and sampling frame refer to the Wave 1 Data Users Guide.
[51] As we do not know the population of SAs who move out of the household, PA survey weights are derived for the original PAs and SA survey weights are derived for the original SAs (even in those cases where SAs moved out of the original household after Wave 1).

Table D.1: Results of multivariate logistic regression modelling PA survey participation in Wave 2

SDB characteristic	Odds ratio	Standard error	p-value
Gender			
Male	1.018	0.099	0.851
Female	1 (Base)		
Site			
Site 1	0.721	0.183	0.199
Site 2	0.264	0.066	0.000***
Site 3	1.318	0.449	0.416
Site 4	0.428	0.178	0.041*
Site 5	0.499	0.169	0.041*
Site 6	0.543	0.125	0.008**
Site 7	0.550	0.195	0.092
Site 8	0.259	0.070	0.000***
Site 9	0.273	0.063	0.000***
Site 10	0.460	0.161	0.027*
Site 11	1 (Base)		
Visa subclass			
200	1 (Base)		
201	1.384	0.646	0.487
202	0.780	0.179	0.282
203	-	-	-
204	0.914	0.138	0.552
866 non-UMA	0.413	0.081	0.000***
866 UMA	1.373	0.222	0.050*
Family size			
0	1 (Base)		
1	1.156	0.186	0.368
2	1.336	0.270	0.153
3	1.775	0.359	0.005**
4 or more	1.769	0.323	0.002***
Country of birth			
Afghanistan	1 (Base)		
Iran	1.036	0.147	0.798
Iraq	1.663	0.198	0.000***
Burma/Myanmar	0.327	0.048	0.000***
Rest of mid-East	1.216	0.219	0.277
Rest of Asia	0.743	0.117	0.060
Africa	0.414	0.065	0.000***
Rest of world	0.689	0.535	0.632
Constant	0.960	0.289	0.893

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Gender included as a control variable.
Locations of sites have not been named to protect participant confidentiality.

Compared to the eligible PA sample, the recruited Wave 2 sample generally had lower representation of PAs living in capital cities. Smaller sized families, participants with a non-UMA visa subclass,[52] and those born in Burma/Myanmar or Africa were also factors associated with a **lower** probability of participation in Wave 2. Those PAs who held a 866 UMA visa subclass and being born in Iraq were significantly associated with a **higher** probability of participation in the second wave.

D.1.2 SA participation in Wave 2

The multivariate logistic model results predicting Wave 2 SA participation are described in Table D.1, which has a breakdown of potential BNLA participants by characteristics drawn from the SDB.

Analysis of factors associated with SA participation in Wave 2 revealed the following participant characteristics were associated with Wave 2 survey completion:

- visa subclass;
- region;
- age;
- country of birth; and
- household structure.

When compared to the eligible SA population, the following characteristics were significantly associated with lower levels of Wave 2 survey completion: older age, living in a capital city, holding an 866 non-UMA visa subclass, being born in Burma, rest of the mid-East or Africa. SA participants who lived in MUs with a male PA and were married were found to have higher levels of survey participation.

[52] A quota was in place for participants holding a non-UMA visa subclass.

Table D.2: Results of multivariate logistic regression modelling SA survey participation in Wave 2

SDB characteristic	Odds ratio	Standard error	p-value
Gender			
Male	0.846	0.100	0.161
Female	1 (Base)		
Site			
Site 1	0.831	0.342	0.655
Site 2	0.224	0.089	0.000***
Site 3	0.865	0.438	0.776
Site 4	1.282	1.081	0.768
Site 5	0.343	0.161	0.023*
Site 6	0.343	0.128	0.004**
Site 7	0.126	0.156	0.096
Site 8	0.205	0.101	0.001**
Site 9	0.291	0.110	0.001**
Site 10	0.094	0.054	0.000***
Site 11	1 (Base)		
Visa subclass			
200	1 (Base)		
201	0.899	0.752	0.900
202	0.869	0.279	0.663
203	-	-	-
204	1.200	0.306	0.475
866 non-UMA	0.327	0.105	0.001**
866 UMA	1.155	0.395	0.674
Age			
Age (continuous years)	0.989	0.003	0.001**
Country of birth			
Afghanistan	1 (Base)		
Iran	1.110	0.236	0.622
Iraq	1.177	0.263	0.465
Burma/Myanmar	0.182	0.054	0.000***
Rest of mid-East	0.416	0.172	0.034*
Rest of Asia	0.685	0.173	0.135
Africa	0.283	0.081	0.000***
Rest of world	-	-	-
Household type			
Male PA	1 (Base)		
Male PA / married female SA	1.620	0.301	0.009**
Male PA / unmarried male or female SA	0.843	0.161	0.374
Male PA / Child 15–17 years	0.712	0.160	0.132
Constant	1.406	0.604	0.428

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Gender included as a control variable.
Locations of sites have not been named to protect participant confidentiality.

D.1.3 Outcomes of Wave 2 survey weighting

The effects of survey weights were examined by comparing unweighted and weighted estimates of the Wave 2 BNLA data with SDB variables not used in the weighting procedure (gender, age group, marital status and migration pathway were considered in this analysis). As Table D.3 shows, in general, the weighting procedure slightly improved the estimation of these proportions for PAs, bringing the BNLA population estimates closer to the SDB total population proportions.

The largest improvement in estimation was found for migration pathway where the survey weighted estimate of 71.3% arriving through an offshore migration pathway was closer to the SDB proportion of 69.5% when compared with the unweighted proportion of 78.1%. Slight improvements were also found in the estimates for age and marital status, when the weighted and unweighted estimates were compared to SDB population totals. However, the estimate of gender was very slightly less accurate when weighted.

Table D.3: Effect of survey weighting on estimates of variables in the Wave 2 BNLA data and SDB population sample

	PAs		
	BNLA W2 unweighted	**BNLA W2 weighted**	**SDB sample**
Gender			
Male	71.1%	71.7%	69.7%
Female	28.9%	28.3%	30.3%
Age group[(a)]			
18–30 years	29.6%	31.9%	32.3%
31–55 years	58.3%	56.1%	55.1%
56 years or over	12.0%	11.9%	12.1%
Marital status			
Married	58.4%	55.3%	55.5%
Migration pathway			
Offshore	78.1%	71.3%	69.5%
Onshore	21.9%	28.7%	30.5%
Total	**1,268**	**1,268**	**4,035**

Note: Information on age group was missing for 0.5% of records in the SDB sample.

D.2 Calculation of Wave 2 longitudinal weights

A set of 'longitudinal' weights that adjusts for attrition between the first and second wave of BNLA data were also calculated and are available in the Wave 2 BNLA dataset. The same process to derive cross-sectional weights described earlier was followed to calculate longitudinal weights. In this case, however, the BNLA estimates were adjusted to the Wave 1 recruited sample (*n* = 1,509 PA, *n* = 890 SA Wave 1 participants).

D.2.1 Factors predicting PAs from Wave 1, participation in Wave 2

The analysis undertaken to generate the BNLA cross-sectional weights was limited to the basic demographic/administrative variables available in the SDB. The analysis used to inform the calculation of longitudinal weights was based on the final Wave 1 and Wave 2 datasets. As comprehensive data was collected on housing, English language proficiency, education and training, employment and income, health, self-sufficiency, community support and perceptions of life in Australia as part of the BNLA study data collection, a much wider range of variables (i.e. those variables that were collected in both Wave 1 and Wave 2 of the BNLA) could be considered in the following analysis. The following factors were tested to determine if they were significantly associated with PAs from Wave 1 participating in the second wave of data collection:

- visa subclass, gender and age;
- capital city;
- household type;
- MU size;
- marital status;
- pre-migration education;
- SEIFA and remoteness index;
- country of birth;
- English language proficiency;
- employment;
- mental health and posttraumatic stress;
- whether waiting for family to migrate to Australia;
- number of times moved home;
- financial hardship and main source of income; and
- physical health.

The multivariate logistic model results predicting PAs from Wave 1 who also participated in Wave 2 are reported in Table D.4.

Table D.4: Results of multivariate logistic regression modelling Wave 1 PAs survey participation in Wave 2

BNLA Wave 1 characteristic	Odds ratio	Standard error	p-value
Country of birth			
Afghanistan	1 (Base)		
Iran	0.699	0.183	0.172
Iraq	1.477	0.332	0.083
Myanmar	1.200	0.386	0.571
Rest of mid-East	1.350	0.590	0.493
Rest of Asia	4.184	1.659	0.000***
Africa	0.747	0.287	0.449
Other confidentialised country	0.281	0.133	0.007**
Pre-migration education			
Never attended school	1 (Base)		
6 or less years of schooling	1.637	0.376	0.032*
7–9 years of schooling	2.192	0.610	0.005**
10–11 years of schooling	1.773	0.578	0.079
12 or more years of schooling	1.631	0.434	0.066
Trade or technical qualifications beyond school	2.278	0.858	0.029*
University Degree	1.436	0.427	0.224
Don't know/prefer not to say	1.630	1.293	0.538
MU structure			
Couple family, with children < 18 years (no other family members)	1 (Base)		
Couple family, with children < 18 years (and other family members)	1.426	0.484	0.295
Couple family only (no other family members)	0.569	0.176	0.070
Couple family and other family members, but no children < 18 years	1.267	0.564	0.595
Single parent family, with children < 18 years (no other family members)	0.910	0.282	0.763
Single parent family, with children < 18 years (and other family members)	2.012	0.820	0.087
Other immediate family members	0.842	0.292	0.621
Single person	0.938	0.187	0.749
Remoteness Area Index (ABS 2011)			
Major Cities of Australia	1 (Base)		
Inner Regional Australia	0.535	0.138	0.016*
Outer Regional Australia	1.836	1.163	0.337
Constant	3.093	0.736	0.000***

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Based on the model results, statistically significant variables predicting PA participation in Wave 2 included:

- country of birth;
- pre-migration education;
- MU structure[53]; and
- remoteness area.

Compared to the eligible Wave 1 PA sample, the recruited Wave 2 PA sample had a higher representation of people born in Iraq and the rest of Asia. Those with more years of pre-migration schooling or a trade or technical qualification were also more likely to participate in Wave 2 compared to participants who had never attended school. PAs living in single parent families with children under 18 years old were also associated with a higher probability of participation.

Conversely, PAs who lived in a couple family only (no other children or family members), and who lived in Inner Regional Australia were both factors associated with lower participation in Wave 2.

D.2.2 Factors predicting Wave 1 SAs participation in Wave 2

Analysis was also undertaken to identify the factors associated with SA applicants from Wave 1 participating in Wave 2. The same factors considered for inclusion in the PA level model, discussed in the previous section were tested in order to identify statistically significant associations with SA participation in Wave 2. An additional factor was also tested, whether the SA was living in a household with a male PA. Results from this modelling are shown in Table D.5 and show that country of birth and English language speaking proficiency were associated with Wave 2 survey completion for SA participants who had completed a Wave 1 survey.

When compared to the eligible Wave 1 SA recruited sample, being born in Iraq or the rest of Asia was associated with a higher probability of survey participation. After controlling for country of birth, participants who had lower levels of English language proficiency were also more likely to participate in the second wave.

[53] Couple family only and single parent family with children under 18 years significant at 0.1 level.

Table D.5: Results of multivariate logistic regression modelling Wave 1 SAs survey participation in Wave 2

BNLA Wave 1 characteristic	Odds ratio	Standard error	p-value
Country of birth			
Afghanistan	1 (Base)		
Iran	1.334	0.450	0.393
Iraq	2.746	0.882	0.002**
Burma/Myanmar	1.431	0.727	0.480
Rest of mid-East	0.981	0.540	0.974
Rest of Asia	4.665	2.89	0.013*
Africa	1.510	0.895	0.487
Other confidentialised country	0.173	0.177	0.087
English language proficiency			
Very well	1 (Base)		
Well	2.976	1.485	0.029*
Not well	2.074	1.003	0.132
Not at all	1.662	0.833	0.311
Don't know/prefer not to say	2.014	1.955	0.471
Constant	1.408	0.800	0.547

Note: * p < 0.05, ** p < 0.01, *** p < 0.001.

D.2.3 Calculation of Wave 2 longitudinal survey weights

As with the calculation of cross sectional survey weights, the appropriateness of generating a single longitudinal weight across both PA and SA participants was further investigated within a variance-component modelling framework. The variance component model of Wave 2 participation of the recruited Wave 1 sample showed that around 1% of the variance in survey completion was explained at the individual level (i.e. PAs and SAs within an MU).[54] A further model with covariates that were found to be statistically significant in the earlier PA and SA level analysis was also estimated. The inclusion of these additional covariates did not explain much further variation in survey participation, with the estimate of the variance of survey completion, explained at the individual level increasing to 5.6% after these covariates were added to the model.

These results indicate that Wave 2 survey participation of Wave 1 recruited participants is largely explained at the MU level. The model estimates indicated that a low level of the variance of survey completion was explained at the individual level, and therefore it was considered appropriate to calculate a survey weight across both PA and SA participants.

The raking procedure described earlier was used to adjust for the differential pattern in non-response of the recruited Wave 1 sample in completing a Wave 2 survey. Longitudinal survey weights were calculated adjusting the Wave 2 estimates to the achieved Wave 1 achieved sample. These survey weights are further described in the next section.

[54] The estimated intra-class correlation was $\boldsymbol{\rho} = 0.989$. This represents that 98.9% of the variance in survey completion was explained at the MU level and 1.1% at the individual level (1 – 0.989 = 0.011 = 1.1% variance).

D.3 Survey weights available in Wave 2 dataset

In total, six survey weights are available in the Wave 2 BNLA dataset. Three 'population' and three 'longitudinal' survey weights are provided in the dataset. These survey weights and their variable names are described further below:

Population weights

1. ***bpawgt*** – PA survey weight. The following variables were used to develop survey weights for this variable: visa subclass, capital city, family size and country of birth. The calculated survey weights ranged in value from 0.40 to 4.25. After an examination of the distribution and factors associated with the larger value survey weights, the calculated survey weights were capped at a maximum value of 2.5 (19 out of 1,268 PAs in Wave 2 had an initial survey weight greater than 2.5). It is recommended that this survey weight be used for cross-sectional analysis involving questions only asked of PAs in Wave 2.
2. ***bsawgt*** – SA survey weight. The visa subclass, capital city, age, country of birth and household structure variables were used to develop survey weights for this variable. SA survey weights ranged from 0.28 to 5.93. As with the PA level weight, the SA level weight was constrained to take a maximum value of 2.5 (involving 13 out of 741 cases). It is recommended that this survey weight be used for cross-sectional analysis involving questions asked only of SAs in Wave 2.
3. ***bwgt*** – all participant survey weight. Visa subclass, capital city, age and country of birth information were used in the survey weighting for this variable. This survey weight ranged from 0.35 to 4.69 and was also capped at a maximum value of 2.5 (involving 25 out of 2,009 cases). It is recommended that this survey weight be used when analysing survey questions of both PA and SA respondents in Wave 2.

Longitudinal weights

1. ***abpawgt*** – PA longitudinal survey weight. The following variables were used to develop survey weights for this variable: country of birth, pre-migration education, MU structure and remoteness index. The calculated survey weights ranged in value from 0.81 to 1.57. It is recommended that this survey weight be used for longitudinal analysis involving questions only asked of PAs in Waves 1 and 2.
2. ***absawgt*** – SA longitudinal survey weight. The country of birth and English speaking proficiency variables were used to develop survey weights for this variable. SA survey weights ranged from 0.87 to 2.64. It is recommended that this survey weight be used for longitudinal analysis involving questions asked only of SAs in Waves 1 and 2.
3. ***abwgt*** – all participant longitudinal survey weight. Country of birth, pre-migration education, MU structure and remoteness information were used in the survey weighting for this variable. This survey weight ranged from 0.82 to 1.42. It is recommended that this survey weight be used when analysing survey questions asked of both PA and SA respondents in Waves 1 and 2.

Appendix E: Wave 3 survey weighting

This section describes the process used to calculate survey weights for the Wave 3 sample of BNLA. The same processes were used to calculate Wave 3 survey weights as was followed for previous waves (See Appendix C and D for further information on the Wave 1 and Wave 2 survey weighting approach).

Two types of survey weights were calculated and are available on the Wave 3 dataset:

1. A population weight, which adjusts BNLA Wave 3 estimates to population totals (in this case, the population of humanitarian migrants who have been granted a permanent visa or arrived in Australia between May and December 2013).[55]
2. A longitudinal weight, which adjusts for attrition between the first and third wave, and second and third wave of BNLA.

E.1 Calculation of Wave 3 population weights

As in previous waves, non-response analysis was undertaken to identify the characteristics of those who completed a Wave 3 interview and how these compared with all PAs and SAs on the DIBP sample frame who were in scope at the time of study recruitment.[56] This non-response analysis was done using multivariate logistic models to identify the factors associated with Wave 3 survey completion. As in Waves 1 and 2, two separate statistical models were estimated—the first identified the factors significantly associated with participation by PA participants and the second model identified factors associated with survey completion by SAs. A third model was also fitted which examined the characteristics associated with survey completion of all participants.

[55] Offshore visa holders had to have **arrived** in Australia holding a permanent visa, while onshore visa holders had to have **received** their protection visa in this time period to be eligible for participation in the BNLA.
[56] *n* = 7,362 individuals in 4,035 MUs were eligible for selection. For more information about the sampling and sampling frame refer to the Wave 1 data users guide.

E.1.1 PA participation in Wave 3

The statistical model results predicting PA participation in a Wave 3 survey are described in Table E.1. The model results show that statistically significant variables predicting PA survey completion, included:

- visa subclass;
- site;
- family size; and
- country of birth.

Compared to the eligible PA sample, the recruited Wave 3 sample generally had lower representation of PAs living in capital cities. Family size was also associated with survey participation, with larger sized families more likely to participate in Wave 3.

Participants with a non-UMA visa subclass[57], and those born in Burma, the rest of Asia or Africa, were also factors associated with a lower probability of participation in Wave 3, while those born in Iraq were more likely to participate in Wave 3 compared to the reference category (those born in Afghanistan).

E.1.2 SA participation in Wave 3

The statistical model results predicting SA participation in a Wave 3 survey are described in Table E.2. Factors associated with SA participation in Wave 3 included the following participant characteristics:

- visa subclass;
- site;
- country of birth; and
- household structure.

When compared to the eligible SA population, the following characteristics were significantly associated with **lower** levels of Wave 3 survey completion: living in a capital city, holding an 866 non-UMA visa subclass, being born in Burma, the rest of Asia, rest of mid-East or Africa. Conversely, SA participants who lived in MUs with a male PA and married female SA were found to have **higher** levels of survey participation.

[57] A quota was in place for participants holding a non-UMA visa subclass.

Table E.1: Results of multivariate logistic regression modelling PA participation in Wave 3

SDB characteristic	Odds ratio	Standard error	p-value
Gender			
Male	1.031	0.104	0.759
Female	1 (Base)		
Site			
Site 1	0.583	0.151	0.038*
Site 2	0.287	0.073	0.000***
Site 3	0.965	0.335	0.920
Site 4	0.223	0.107	0.002**
Site 5	0.689	0.234	0.275
Site 6	0.430	0.101	0.000***
Site 7	0.346	0.131	0.005**
Site 8	0.323	0.088	0.000***
Site 9	0.226	0.062	0.000***
Site 10	0.679	0.236	0.268
Site 11	1 (Base)		
Visa subclass			
200	1 (Base)		
201	1.194	0.590	0.718
202	0.633	0.155	0.063
203	-	-	-
204	1.001	0.153	0.993
866 non-UMA	0.356	0.077	0.000***
866 UMA	1.075	0.185	0.674
Family size			
0	1.036	0.181	0.836
1	1 (Base)		
2	1.375	0.192	0.023*
3	1.824	0.247	0.000***
4 or more	1.783	0.180	0.000***
Country of birth			
Afghanistan	1 (Base)		
Iran	0.846	0.124	0.258
Iraq	1.492	0.180	0.001**
Burma/Myanmar	0.291	0.044	0.000***
Rest of mid-East	0.981	0.189	0.922
Rest of Asia	0.618	0.102	0.004**
Africa	0.324	0.054	0.000***
Rest of world	0.373	0.390	0.347
Constant	1.081	0.284	0.766

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Gender included as a control variable.
Locations of sites have not been named to protect participant confidentiality.

Table E.2: Results of multivariate logistic regression modelling SA participation in Wave 3

SDB characteristic	Odds ratio	Standard error	p-value
Gender			
Male	0.846	0.101	0.165
Female	1 (Base)		
State			
Site 1	0.697	0.297	0.399
Site 2	0.311	0.127	0.004**
Site 3	0.550	0.313	0.295
Site 4	1.729	1.476	0.521
Site 5	0.497	0.232	0.136
Site 6	0.329	0.129	0.005**
Site 7	0.125	0.161	0.107
Site 8	0.237	0.118	0.004**
Site 9	0.271	0.107	0.001**
Site 10	0.104	0.060	0.000***
Site 11	1 (Base)		
Visa subclass			
200	1 (Base)		
201	0.469	0.553	0.521
202	0.871	0.304	0.694
203	-	-	-
204	1.329	0.340	0.266
866 non-UMA	0.256	0.095	0.000***
866 UMA	0.614	0.236	0.205
Country of birth			
Afghanistan	1 (Base)		
Iran	0.903	0.192	0.633
Iraq	0.862	0.190	0.501
Burma/Myanmar	0.138	0.040	0.000***
Rest of mid-East	0.315	0.124	0.004**
Rest of Asia	0.482	0.121	0.004**
Africa	0.145	0.039	0.000***
Rest of world	0.979	1.084	0.985
Household type			
Male PA	1 (Base)		
Male PA / married female SA	1.898	0.350	0.001**
Male PA / unmarried male or female SA	1.001	0.194	0.995
Male PA / Child 15–17 years	1.032	0.231	0.888
Constant	1.183	0.517	0.701

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Gender included as a control variable.
Locations of sites have not been named to protect participant confidentiality.

E.1.3 Calculation of Wave 3 cross-sectional survey weights

Further analysis was undertaken to identify the appropriateness of calculating a single survey weight across PA and SA respondents using variance component modelling. The rationale for this approach and further description of the variance component modelling is described in Appendix D.

Estimates from the unadjusted variance component modelling showed that 8.9% of the variance in survey completion was explained at the individual level (i.e. PAs and SAs within an MU).[58] A second model with explanatory covariates was also estimated.[59] The inclusion of these additional variables added little explanatory power when these results were compared to the unadjusted model. The estimate of variance in Wave 3 survey completion at the individual level was 14.1%. Given the low level of variance explained at the individual level and that survey completion was largely driven by variance at the MU level, a single Wave 3 survey weight was also calculated, and this is described next.

To adjust for the differential pattern in non-response described above, Wave 3 survey weights were calculated using the *ipfweight* algorithm (also known as raking) in Stata authored by Michael Bergmann.[60] The *ipfweight* procedure is based on work first proposed by Deming and Stephan (1940)[61] and adjusts survey sample weights to achieve population totals. In this case the population of humanitarian migrants being granted a permanent visa between May and December 2013.[62]

E.1.4 Outcomes of Wave 3 survey weighting

The effects of survey weighting were further explored by comparing unweighted and weighted estimates of the Wave 3 BNLA data with SDB variables not used in the weighting procedure (gender, age group, marital status and migration pathway were considered in this analysis). As Table E.3 shows, in general, there was a slight improvement in the accuracy of population estimates through the survey weighting process, with survey weights bringing the BNLA population estimates closer to the SDB total population proportions.

The largest improvement in estimation was found for migration pathway where the survey weighted estimate of 72.2% arriving through an offshore migration pathway was closer to the SDB proportion of 69.5% when compared with the unweighted proportion of 80.8%. There were also slight improvements found for the estimates of age and marital status. However, as in previous waves, the estimate for gender was slightly less accurate when the survey weighted estimated was compared to the unweighted estimated.

[58] The estimated intra-class correlation was **ρ** = 0.912. This represents that 91.2% of the variance in survey completion was explained at the MU level and 8.9% at the individual level (1 – 0.912 = .088 = 8.9% variance).
[59] Gender, visa subclass, site, family size and country of birth were found to be significantly associated with Wave 3 survey completion and included in the variance-components model.
[60] See http://fmwww.bc.edu/repec/bocode/i/ipfweight.html for further detail. Accessed on 29 April 2014.
[61] Deming, W. Edwards/Frederick F. Stephan. 1940. On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals Are Known, in: *The Annals of Mathematical Statistics* 11 (4): 427-444.
[62] Offshore visa holders had to have **arrived** in Australia holding a permanent visa, while onshore visa holders had to have **received** their protection visa in this time period.

Table E.3: Effect of survey weighting on estimates of variables in the Wave 3 BNLA data and SDB population sample for PAs

	BNLA Wave 3 unweighted	BNLA Wave 3 weighted	SDB Sample
Gender			
Male	69.8%	70.9%	69.7%
Female	30.2%	29.1%	30.3%
Age group[(a)]			
18–30 years	26.9%	29.9%	32.8%
31–55 years	59.8%	57.4%	55.1%
56 years or over	13.3%	12.7%	12.1%
Marital status			
Married	60.1%	56.6%	55.5%
Migration pathway			
Offshore	80.8%	72.2%	69.5%
Onshore	19.2%	27.8%	30.5%
Total	**1,155**	**1,155**	**4,035**

(a) Information on age group was missing for 0.5% of records in the SDB sample.

E.2 Calculation of Wave 3 longitudinal weights

A set of longitudinal weights that adjusts for attrition between the first and third wave of BNLA, and the second and third waves were also calculated and are available in the Wave 3 BNLA dataset. The same process to derive cross-sectional weights described earlier was followed to calculate longitudinal weights. In this case, however, the BNLA estimates were adjusted to the Wave 1 recruited sample (*n* =1,509 PA, *n* =890 SA Wave 1 participants) to calculate the Wave 1/3 longitudinal weight. For the Wave 2/3 longitudinal weight, the Wave 3 estimates were adjusted to the Wave 2 recruited sample (*n* =1,268 PA, *n* =741 SA participants in Wave 2).

E.2.3 Factors predicting PAs' from Wave 1 participation in Wave 3

The analysis undertaken to generate the BNLA cross-sectional weights was limited to the basic demographic/administrative variables available in the SDB. The analysis used to inform the calculation of longitudinal weights was based on the final Wave 1 and Wave 3 datasets. As comprehensive data was collected on housing, English language proficiency, education and training, employment and income, health, self-sufficiency, community support and perceptions of life in Australia as part of the BNLA study data collection, a much wider range of variables (i.e. those variables that were collected in both Wave 1 and Wave 3 of the BNLA) could be considered in the following analysis.

The following factors were tested to determine if they were significantly associated with PAs from Wave 1 participating in the second wave of data collection:

- visa subclass, gender and age;
- capital city;
- household type;
- MU size;
- marital status;
- pre-migration education;
- SEIFA and remoteness index;
- country of birth;
- English language proficiency;
- employment;
- mental health and posttraumatic stress;
- whether waiting for family to migrate to Australia;
- number of times moved home;
- financial hardship and main source of income; and
- physical health.

The multivariate logistic model results predicting PAs from Wave 1 who also participated in Wave 3 are described in Table E.4.

Table E.4: Results of multivariate logistic regression modelling Wave 1 PAs survey participation in Wave 3

BNLA Wave 1 characteristic	Odds ratio	Standard error	p-value
Age			
Age (years)	1.021	0.006	0.001**
Visa subclass			
200	1 (Base)		
201	0.669	0.597	0.653
202	0.484	0.186	0.060
203	-	-	-
204	0.979	0.248	0.936
866 non-UMA	0.532	0.164	0.041*
866 UMA	0.807	0.155	0.267
Country of birth			
Afghanistan	1 (Base)		
Iran	0.562	0.133	0.015*
Iraq	1.115	0.238	0.608
Myanmar	0.852	0.249	0.585
Rest of mid-East	0.904	0.368	0.806
Rest of Asia	1.139	0.289	0.608
Africa	0.445	0.155	0.021*
Other confidentialised country	0.354	0..175	0.036*
Pre-migration education			
Never attended school	1 (Base)		
6 or less years of schooling	2.148	0.490	0.001**
7–9 years of schooling	1.715	0.431	0.032*
10–11 years of schooling	1.599	0.461	0.104
12 or more years of schooling	2.192	0.588	0.003**
Trade or technical qualifications beyond school	1.657	0.552	0.129
University Degree	1.615	0.491	0.113
Don't know/prefer not to say	2.244	1.553	0.243
MU structure			
Couple family, with children < 18 years (no other family members)	1 (Base)		
Couple family, with children < 18 years (and other family members)	1.367	0.430	0.320
Couple family only (no other family members)	0.742	0.222	0.321
Couple family and other family members, but no children < 18 years	1.650	0.842	0.326
Single parent family, with children < 18 years (no other family members)	1.025	0.324	0.938
Single parent family, with children < 18 years (and other family members)	3.798	2.004	0.011*
Other immediate family members	0.789	0.273	0.495
Single person	0.761	0.138	0.134
Understanding of spoken English			
Very well	1 (Base)		
Well	0.536	0.149	0.025*
Not well	0.711	0.208	0.245
Not at all	0.807	0.278	0.535
Don't know/prefer not to say	0.491	0.322	0.279
Constant	1.709	0.713	0.199

Note: * p < 0.05, ** p < 0.01, *** p < 0.001.

Based on the model results, statistically significant variables predicting PA participation in Wave 3 included:

- age;
- visa subclass;
- country of birth;
- pre-migration education;
- MU structure; and
- English language proficiency.

Compared to the eligible Wave 1 PA sample, the recruited Wave 3 PA sample had a lower representation of participants who were born in Iran, Africa or other remaining confidentialised countries. PA participants holding an 866 non-UMA visa type and poorer language proficiency were less likely to complete a Wave 3 interview.

Particular groups of PAs were more likely to complete a Wave 3 interview. These included: older PAs; those with more pre-migration education (compared to those who had never attended school); and those PAs who migrated to Australia as a single parent family with children under 18 years old (compared to those who migrated as a couple family with children under 18 years).

E.2.2 Factors predicting SAs' from Wave 1 participation in Wave 3

Analysis was also undertaken to identify the factors associated with SAs' from Wave 1 participation in Wave 3. The same factors considered for inclusion in the PA level model discussed in the previous section were tested in order to identify statistically significant associations with SA participation in Wave 3. An additional factor was also tested, whether the SA was living in a household with a male PA. Results from this modelling are shown in Table E.5 and show that age, visa subclass, country of birth, pre-migration education levels and MU structure were associated with Wave 3 survey completion for SA participants who had previously completed a Wave 1 survey.

When compared to the eligible Wave 1 SA recruited sample, being older was associated with a higher probability of Wave 3 survey participation. SA participants born in rest of mid-East, Africa or other remaining confidentialised countries were less likely to participate in Wave 3.

Table E.5: Results of multivariate logistic regression modelling Wave 1 SAs survey participation in Wave 3

BNLA Wave 1 characteristic	Odds ratio	Standard error	p-value
Age			
Age (years)	1.020	0.009	0.032*
Visa subclass			
200	1 (Base)		
201	1.070	1.831	0.968
202	0.655	0.418	0.508
203	-	-	-
204	1.546	0.705	0.338
866 non-UMA	0.607	0.521	0.562
866 UMA	0.164	0.098	0.003**
Country of birth			
Afghanistan	1 (Base)		
Iran	0.730	0.300	0.445
Iraq	0.920	0.337	0.840
Myanmar	0.489	0.295	0.236
Rest of mid-East	0.179	0.128	0.016*
Rest of Asia	1.827	1.238	0.373
Africa	0.293	0.163	0.028*
Other confidentialised country	0.121	0.101	0.011*
Pre-migration education			
Never attended school	1 (Base)		
6 or less years of schooling	1.360	0.582	0.472
7–9 years of schooling	0.890	0.377	0.785
10–11 years of schooling	0.966	0.431	0.939
12 or more years of schooling	0.661	0.271	0.313
Trade or technical qualifications beyond school	0.309	0.170	0.033*
University Degree	1.198	0.635	0.732
Don't know/prefer not to say	0.662	0.612	0.656
MU structure			
Couple family, with children < 18 years (no other family members)	1 (Base)		
Couple family, with children < 18 years (and other family members)	0.608	0.206	0.144
Couple family only (no other family members)	0.399	0.151	0.016*
Couple family and other family members, but no children < 18 years	1.535	0.777	0.397
Single parent family, with children < 18 years (no other family members)	1.311	1.157	0.758
Single parent family, with children < 18 years (and other family members)	0.609	0.263	0.253
Other immediate family members	0.404	0.211	0.084
Constant	5.960	3.354	0.002**

Note: * p < 0.05, ** p < 0.01, *** p < 0.001.

E.2.3 Factors predicting participants from Wave 2 completing a Wave 3 interview

The same process described in the previous section was used to identify the factors associated with completion of a Wave 3 interview for those who had previously participated in Wave 2. As before, separate models were estimated to identify the factors associated with PA and SA participation. In total, 1,704 participants who completed a Wave 2 interview also completed a Wave 3 interview.[63] The same factors that were tested in the model predicting survey participation in Wave 3 for those who completed a Wave 1 interview were also tested in these models. The model results for PAs are shown in Table E.6, with results for SA reported in Table E.7.

PAs who completed a Wave 2 interview were more likely to participate in Wave 3 if they were older and lived in more advantaged areas at Wave 2 (based on SEIFA 2011 relative to those living in the bottom decile).

Compared to the eligible Wave 2 PA sample, the recruited Wave 3 PA sample had a lower representation of participants who were born in Iran, or Africa. PA participants holding an 866 non-UMA or 866 UMA visa type were less likely to complete a Wave 3 interview.

Looking now at the factors associated with SA participation in Wave 3, for those SAs who had previously completed a Wave 2 interview, model results for SA participants are presented in Table E.7.

SAs who participated in Wave 2 and held a 202 or 866 UMA visa type were less likely to participate in Wave 3 compared to those holding a 200 visa. As were those whose highest level of pre-migration education was a trade or technical certificate, and participants who migrated to Australia as a couple family with children under 18 years old. Participants who reported a greater number of financial hardship events also were less likely to participate in Wave 3.

English language proficiency was also associated with higher levels of survey participation in Wave 3. SA participants who had lower levels of understanding of spoken English were less likely to complete a Wave 3 interview compared to those who understood spoken English "very well".

[63] 190 participants who completed a Wave 1 interview, did not participate in Wave 2 but returned to complete an interview at Wave 3.

Table E.6: Results of multivariate logistic regression modelling Wave 2 PAs' survey participation in Wave 3

BNLA Wave 2 characteristic	Odds ratio	Standard error	p-value
Age			
Age (years)	1.034	0.007	0.000***
Visa subclass			
200	1 (Base)		
201	0.334	0.318	0.250
202	0.498	0.229	0.131
203	-	-	-
204	0.917	0.246	0.747
866 non-UMA	0.489	0.165	0.035*
866 UMA	0.646	0.142	0.048*
Country of birth			
Afghanistan	1 (Base)		
Iran	0.625	0.174	0.093
Iraq	1.399	0.322	0.146
Myanmar	0.869	0.297	0.682
Rest of mid-East	1.065	0.508	0.895
Rest of Asia	0.885	0.235	0.648
Africa	0.365	0.145	0.012*
Other confidentialised country	0.395	0.256	0.153
SEIFA 2011 decile			
1	1 (Base)		
2	2.080	0.516	0.003**
3	1.775	0.488	0.037*
4	0.857	0.313	0.673
5	1.079	0.309	0.789
6	1.733	0.665	0.152
7	0.892	0.318	0.750
8	1.902	0.992	0.218
9	2.325	1.334	0.141
10	0.897	0.427	0.820
Constant	1.327	0.438	0.392

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table E.7: Results of multivariate logistic regression modelling Wave 2 SAs' survey participation in Wave 3

BNLA Wave 2 characteristic	Odds ratio	Standard error	p-value
Visa subclass			
200	1 (Base)		
201	0.218	0.294	0.259
202	0.262	0.172	0.041*
203	-	-	-
204	0.756	0.504	0.676
866 non-UMA	0.384	0.293	0.211
866 UMA	0.099	0.063	0.000***
Pre-migration education			
Never attended school	1 (Base)		
6 or less years of schooling	1.027	0.575	0.961
7–9 years of schooling	0.855	0.484	0.783
10–11 years of schooling	0.930	0.553	0.904
12 or more years of schooling	0.522	0.289	0.242
Trade or technical qualifications beyond school	0.278	0.178	0.047*
University Degree	0.812	0.542	0.755
Don't know/prefer not to say	0.628	0.619	0.637
MU structure (at Wave 1)			
Couple family, with children < 18 years (no other family members)	1 (Base)		
Couple family, with children < 18 years (and other family members)	0.374	0.148	0.013*
Couple family only (no other family members)	0.826	0.497	0.751
Couple family and other family members, but no children < 18 years	1.248	0.729	0.705
Single parent family, with children < 18 years (no other family members)	2.750	3.195	0.384
Single parent family, with children < 18 years (and other family members)	0.547	0.391	0.400
Other immediate family members	0.871	0.583	0.837
Understanding of spoken English			
Very well	1 (Base)		
Well	2.666	1.229	0.033*
Not well	2.956	1.454	0.028*
Not at all	5.370	3.305	0.006**
Don't know/prefer not to say			
Number of financial hardship events			
Number of financial hardship events	0.988	0.004	0.011*
Constant	6.833	5.020	0.009**

Note: * p < 0.05, ** p < 0.01, *** p < 0.001.

E.2.4 Calculation of Waves 1 and 3, and Waves 2 and 3 longitudinal survey weights

As before, the appropriateness of generating a single longitudinal weight across both PA and SA participants was further investigated within a variance-component modelling framework.

The variance component model of Wave 3 participation of the recruited Wave 1 sample showed that that around 2% of the variance in survey completion was explained at the individual level (i.e. PAs and SAs within an MU). An even smaller proportion of variance (<1%) in Wave 3 survey completion was explained at the individual level for participants who completed a Wave 2 interview. Including additional explanatory variables in both models did not explain much further variance in survey participation: 4.8% of the variance was explained in the Wave 1 and 3 model after additional covariates were added, with the same proportion (4.8%) of variation explained in the Wave 2 and 3 model after additional variables were added to the variance components model.

For the reasons described earlier, a single longitudinal weight was therefore also calculated for all participants for both Waves 1 and 3; and Waves 2 and 3.

The raking procedure described earlier was used to adjust for the differential pattern in non-response of the recruited Wave 1 and Wave 2 sample. Longitudinal weights were calculated that adjusted the Wave 3 estimates to the achieved Wave 1 and Wave 2 samples. These survey weights are further described in the last section of this Appendix.

E.2.5 Factors predicting participation in all 3 waves

This section describes the calculation of the longitudinal survey weight for respondents to the first three waves or 'the balanced panel survey weight'. This weight applies the group of participants who were interviewed in Wave 1 and responded in each subsequent wave up to and including Wave 3. This group is counted 'as responses' and every other participant in Wave 1 who did not complete each subsequent wave as 'non-responses'. The analysis that follows identifies the characteristics associated participants who responded in all 3 waves and these characteristics are used in the later survey weighting procedure to adjust for the differential pattern of responses between those who completed all three waves compared with those who did not complete at least one interview. Multivariate logistic models were estimated to identify the characteristics associated with survey completion. Separate models were estimated for PA, SA and all participants. Overall, 1,053 PAs and 651 SAs completed a survey in all three waves. This means a total of 1,704 participants completed all 3 waves of data collection.

The analysis used to inform the calculation of longitudinal weights was based on the final Waves 1 to 3 datasets. The following factors were tested to determine if they were significantly associated with survey completion across all 3 waves:

- Visa sub-class, gender and age
- Capital city
- Household type
- MU size
- Marital status

- Pre-migration education
- SEIFA and remoteness index
- Country of birth
- English language proficiency
- Employment
- Mental health and posttraumatic stress
- Whether waiting for family to migrate to Australia
- Number of times moved home
- Financial hardship and main source of income
- Physical health

Statistical model results predicting participation in all 3 waves for each participant type are summarised in Table E.8. The model results show that PAs were more likely to participate across all 3 waves of data collection if they had higher levels of schooling (compared to those who had never attended school), migrated to Australia as a single parent with children under 18 years or were classified as 'may have Post Traumatic Stress Disorder' using the PTSD-8 measure collected in the survey. PAs who arrived in Australia with an 866 non-UMA visa subclass, or were born in in Iran, Africa or the rest of the world had lower likelihood of participation in all 3 waves of data collection.

Visa sub-class and country of birth were the only two factors associated with SA survey completion. SA participants holding an 866-UMA visa were less likely to participate in all three waves compared to 200 visa holders. SAs born in Iraq or the rest of Asia were more likely to participate in all three waves relative to those born in Afghanistan.

Table E.8: Results of multivariate logistic regression modelling survey participation in the first three waves, by participant type at Wave 1

BNLA Wave 1 characteristic	PA model (Odds Ratio)	SA model (Odds Ratio)	All participants (Odds Ratio)
Pre-migration education			
Never attended school	1 (Base)		1 (Base)
6 or less years of schooling	1.809**		1.428*
7-9 years of schooling	1.646*		1.566*
10-11 years of schooling	1.435		1.571*
12 or more years of schooling	1.818*		1.296
Trade or technical qualifications beyond school	1.848*		1.281
University Degree	1.368		1.303
Don't know/prefer not to say	2.503		1.504
Age (years)	1.020***		1.016***
Visa Sub-Class			
200	1 (Base)	1 (Base)	1 (Base)
201	0.410	1.023	0.679
202	0.617	0.746	0.615
203	-	-	-
204	0.581*	1.275	0.706
866-non-UMA	0.555*	0.997	0.715
866-UMA	0.842	0.266*	0.625*
Post-traumatic stress (PTSD 8)			
Unlikely to have Post Traumatic Stress Disorder	1 (Base)		
May have Post Traumatic Stress Disorder	1.349*		
Don't know/prefer not to say	0.311***		
Country of birth			
Afghanistan	1 (Base)	1 (Base)	1 (Base)
Iran	0.634*	1.089	0.765
Iraq	1.257	2.094*	1.544*
Burma/Myanmar	1.087	1.524	1.042
Rest of mid-East	1.382	0.506	0.887
Rest of Asia	1.571	4.724**	1.955**
Africa	0.475*	0.806	0.549**
Rest of world	0.306*	0.113	0.207***
Migrating unit structure			
Couple family with children under 18 (no other family members)	1 (Base)		1 (Base)
Couple family with children under 18 and other family members	1.048		0.826
Couple family only (no other family members)	0.638		0.632*
Couple family and other family members but no children under 18	1.122		1.066
Single parent family with children under 18 (no other family members)	1.056		1.102
Single parent family with children under 18 and other family members	2.523*		1.887*
Other immediate family members	1.075		0.958
Other extended family members only	-		-
Non-related persons	-		-
Single person	0.841		0.753*
Constant	**0.817**	**1.929**	**1.119**

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

E.3 Survey weights available in the Wave 3 dataset

In total, twelve survey weights are available on the Wave 3 BNLA dataset. Three population and nine longitudinal survey weights are provided on the dataset. These survey weights and their variable names are described further below:

Population weights

1. ***cpawgt –*** Wave 3 PA Survey weight. The following variables were used to develop survey weights for this variable: visa subclass, capital city, family size and country of birth. The calculated survey weights ranged in value from 0.30 to 5.43. After an examination of the distribution and factors associated with the larger value survey weights, the calculated survey weights were capped at a maximum value of 2.5.

   It is recommended that this survey weight be used for cross-sectional analysis involving questions only asked of PAs in Wave 3.

2. ***csawgt –*** Wave 3 SA survey weight. The visa subclass, capital city, age, country of birth and household structure variables were used to develop survey weights for this variable. SA survey weights ranged from 0.25 to 5.96. As with the PA level weight, the SA level weight was constrained to take a maximum value of 2.5.

   It is recommended that this survey weight be used for cross-sectional analysis involving questions asked only of SAs in Wave 3.

3. ***cwgt*** – Wave 3 all participant survey weight. Visa subclass, capital city, age and country of birth information were used in the survey weighting for this variable. This survey weight ranged from 0.39 to 4.60 and was also capped at a maximum value of 2.5.

   It is recommended that this survey weight be used when analysing survey questions of both PA and SA respondents in Wave 3.

Longitudinal weights

1. ***acpawgt*** – Wave 1/3 PA longitudinal survey weight. The following variables were used to develop survey weights for this variable: visa subclass, age, country of birth, pre-migration education, MU structure and English language proficiency. The calculated survey weights ranged in value from 0.79 to 1.46.

   It is recommended that this survey weight be used for longitudinal analysis involving questions only asked of PAs in Waves 1 and 3.

2. ***acsawgt*** – Wave 1/3 SA longitudinal survey weight. The visa subclass, country of birth, age, pre-migration education and MU structure variables were used to develop survey weights for this variable. Wave 1/3 SA longitudinal survey weights ranged from 0.92 to 1.78.

   It is recommended that this survey weight be used for longitudinal analysis involving questions asked only of SAs in Waves 1 and 3.

3. ***acwgt*** – Wave 1/3 all participant longitudinal survey weight. Visa subclass, country of birth, age, pre-migration education, and MU structure information were used in the survey weighting for this variable. This survey weight ranged from 0.83 to 1.50.

It is recommended that this survey weight be used when analysing survey questions asked of both PA and SA respondents in Waves 1 and 3.

4. ***bcpawgt*** – Wave 2/3 PA longitudinal survey weight. The following variables were used to develop survey weights for this variable: visa subclass, age, country of birth, and SEIFA 2011. The calculated survey weights ranged in value from 0.85 to 1.32.

   It is recommended that this survey weight be used for longitudinal analysis involving questions only asked of PAs in Waves 2 and 3.

5. ***acsawgt*** – Wave 1/3 SA longitudinal survey weight. The visa subclass, pre-migration education, MU structure, understanding of spoken English and number of financial hardship events, variables were used to develop survey weights for this variable. Wave 2/3 SA longitudinal survey weights ranged from 0.86 to 1.57.

   It is recommended that this survey weight be used for longitudinal analysis involving questions asked only of SAs in Waves 2 and 3.

6. ***bcwgt*** – Wave 2/3 all participant longitudinal survey weight. Visa subclass, age, country of birth, and SEIFA 2011 were used in the survey weighting for this variable. This survey weight ranged from 0.87 to 1.32.

   It is recommended that this survey weight be used when analysing survey questions asked of both PA and SA respondents in Waves 2 and 3.

7. ***a_cpawgt*** – Waves 1, 2, and 3 balanced panel longitudinal weight for Principal Applicants. The following variables were used to develop survey weights for this variable: visa sub-class, country of birth, pre-migration education, age, and posttraumatic stress. The calculated survey weights ranged in value from 0.82 to 1.55.

8. ***a_csawgt*** – Waves 1, 2, and 3 balanced panel longitudinal survey weight for Secondary Applicants. The visa sub-class and country of birth variables were used to develop this survey weight. This longitudinal survey weight ranged from 0.88 to 1.72.

9. ***a_cwgt*** – Waves 1, 2, and 3 balanced panel longitudinal survey weight for all participants. Visa sub-class, country of birth, age, and pre-migration education variables were used in the survey weighting for this variable. This survey weight ranged from 0.74 to 1.87.

Appendix F: Wave 4 survey weighting

This section describes the process used to calculate survey weights for the Wave 4 sample of Building a New Life in Australia. The same processes were used to calculate Wave 4 survey weights as for previous waves (See Appendix C, D and E for further information on the BNLA survey weighting approach).

Two types of survey weights were calculated and are available in the Wave 4 dataset:

1. A 'population' weight – that adjusts BNLA Wave 4 estimates to population totals (in this case, the population of humanitarian migrants who have been granted a permanent visa or arrived in Australia between May and December 2013).[64]
2. A 'longitudinal' weight that adjusts for attrition between the first and fourth wave, and a balanced panel survey weight for respondents who completed all of the first four waves of data collection.

F.1 Calculation of Wave 4 population weights

Analysis of non-response was undertaken to identify the characteristics of those who completed an interview compared with all humanitarian migrants on the SDB that were eligible to participate in the study at the time of study recruitment.[65] To identify the factors associated with participation in Wave 4, multivariate logistic models were estimated. As in previous waves, separate models were run to identify the characteristics associated with PA participation, SA participation, and survey completion by all participants.

F.1.1 Factors associated with participation in Wave 4

Results of statistical models predicting participation for each participant type are summarised in Table F.1.

[64] Offshore visa holders had to have **arrived** in Australia holding a permanent visa, while onshore visa holders had to have **received** their protection visa in this time period to be eligible for participation in BNLA.

[65] n=7,362 individuals in 4,035 migrating units were eligible for selection. For more information about the sampling and sampling frame refer to the Sections 2.1 and 2.2 of this Data Users Guide.

Table F.1: Results of multivariate logistic regression predicting survey participation in Wave 4, by participant type, odds ratios

SDB characteristic	PA model	SA model	All participants
Gender			
Male	1.086	0.804	0.594***
Female	1 (Base)	1 (Base)	1 (Base)
Site			
Site 1[#]	0.606	0.649	0.653
Site 2[#]	0.277***	0.287**	0.305***
Site 3	1.045	0.691	0.960
Site 4[#]	0.288**	0.219	0.224**
Site 5[#]	0.744	0.520	0.734
Site 6[#]	0.425***	0.295**	0.376***
Site 7	0.355**	-	0.221***
Site 8[#]	0.303***	0.185**	0.260***
Site 9[#]	0.226***	0.263**	0.245***
Site 10	0.562	0.114***	0.291***
Site 11	1 (Base)	1 (Base)	1 (Base)
Visa Sub-Class			
200	1 (Base)	1 (Base)	1 (Base)
201	1.489	1.209	1.368
202	0.734	0.753	0.831
203	-	-	
204	1.095	1.140	1.113
866-non-UMA	0.413***	0.277*	0.337***
866-UMA	0.928	0.819	0.712**
Family Size			
0	1.022		
1	1 (Base)		
2	1.346*		
3	1.435**		
4 or more	1.761***		
Country of birth			
Afghanistan	1 (Base)	1 (Base)	1 (Base)
Iran	0.890	0.938	0.918
Iraq	1.747***	0.901	1.492**
Burma/Myanmar	0.333***	0.135***	0.275***
Rest of mid-East	0.788	0.418**	0.635*
Rest of Asia	0.626**	0.484**	0.626**
Africa	0.329***	0.137***	0.265***
Rest of world	0.322	0.930	0.439
Household type			
Male PA		1 (Base)	1 (Base)
Male PA / married female SA		1.664**	0.449***
Male PA / unmarried male or female SA		0.987	0.319***
Male PA / Child 15-17 years		0.938	0.300***
Constant	1.113	1.355	3.059***

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Gender included as a control variable.
[#] Indicates site in a capital city. Location of sites have not been named to protect participant confidentiality.

The model results show that both PA and SA participants share some factors that predict survey completion in Wave 4. For both the PA and SA sample, participants living in capital cities and holding an 866 non-UMA[66] visa were significantly less likely to participate in Wave 4. Country of birth was also significantly associated with Wave 4 survey completion for both PAs and SAs, although there were some differences between the two participant types. PAs born in Iraq were more likely to participate in Wave 4 compared to those born in Afghanistan, however no significant association with survey completion was found for SA participants born in Iraq. SAs who were born in the remainder of the Middle East were less likely to complete a Wave 4 survey compared to SAs born in Afghanistan, however this association was not significant for PAs. Both PA and SA participants born in Burma/Myanmar, Rest of Asia and Africa were less likely to participate in Wave 4.

Family size was significantly associated with PA participation in Wave 4. PAs with larger families were more likely to complete a Wave 4 survey. Household type was significantly associated with SA participation but was not found to be significant for PA participants.

F.1.2 Calculation of Wave 4 cross-sectional survey weights

Further analysis was undertaken to identify the appropriateness of calculating a single survey weight across PA and SA respondents using variance component modelling. The rationale for this approach and further description of the variance component modelling is provided in Appendix C.

Estimates from the unadjusted variance component modelling showed that 11.2% of the variance in survey completion was explained at the individual level (i.e. PAs and SAs within a migrating unit).[67] A second model with explanatory covariates was also estimated.[68] The inclusion of these additional variables added little explanatory power when these results were compared to the unadjusted model. The estimate of variance in Wave 4 survey completion at the individual level was 15.2%. Given the low level of variance explained at the individual level and that survey completion was largely driven by variance at the migrating unit level, a single Wave 4 survey weight was also calculated, and this is described next.

To adjust for the differential pattern in non-response described above, Wave 4 survey weights were calculated using the *ipfweight* algorithm (also known as raking) in Stata authored by Michael Bergmann.[69] The *ipfweight* procedure is based on work first proposed by Deming and Stephan (1940)[70] and adjusts survey sample weights to achieve population totals, in this case the population of humanitarian migrants being granted a permanent visa between May and December 2013.[71]

[66] A quota was in place for participants holding a non-UMA visa subclass.
[67] The estimated intra-class correlation was **ρ** = 0.888. This indicates that 88.8% of the variance in survey completion was explained at the migrating unit level and 11.2% at the individual level (1 – 0.888 = 0.112 = 11.2% variance).
[68] Gender, visa sub-class, site, family size and country of birth were found to be significantly associated with Wave 4 survey completion and included in the variance-components model.
[69] See http://fmwww.bc.edu/repec/bocode/i/ipfweight.html for further detail. Accessed on 29/4/2014.
[70] Deming, W. Edwards/Frederick F. Stephan. 1940. On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals Are Known, in: *The Annals of Mathematical Statistics* 11 (4): 427-444.
[71] Offshore visa holders had to have **arrived** in Australia holding a permanent visa, while onshore visa holders had to have **received** their protection visa in this time period.

F.1.3 Outcomes of Wave 4 survey weighting

The effects of survey weighting were further explored by comparing unweighted and weighted estimates of the Wave 4 BNLA data with Settlement Database (SDB) variables not used in the weighting procedure (gender, age group, marital status and migration pathway were considered in this analysis). As Table F.2 shows, in general, there was a slight improvement in the accuracy of population estimates through the survey weighting process, with survey weights bringing the BNLA population estimates closer to the SDB total population proportions.

The largest improvement in estimation was found for migration pathway where the survey weighted estimate of 71.9% arriving through an offshore migration pathway was closer to the SDB proportion of 69.5% when compared with the unweighted proportion of 82.1%. There were also slight improvements found for the estimates of age and marital status. However, as in previous waves, the estimate for gender was slightly less accurate when the survey weighted estimate was compared to the unweighted estimate.

Table F.2: Effect of survey weighting on estimates of variables in the Wave 3 BNLA data and SDB population sample

	Principal Applicants		
	BNLA Wave 4 unweighted	**BNLA Wave 4 weighted**	**SDB Sample**
Gender			
Male	69.8%	71.4%	69.7%
Female	30.2%	28.6%	30.3%
Age Group[(a)]			
18-30 years	27.9%	31.1%	32.8%
31-55 years	59.1%	56.6%	55.1%
56 years or over	13.0%	12.3%	12.1%
Marital status			
Married	59.3%	55.9%	55.5%
Migration pathway			
Offshore	82.1%	71.9%	69.5%
Onshore	17.9%	28.1%	30.5%
Total	**1,186**	**1,186**	**4,035**

Notes: (a) Information on age group was missing for 0.5% of records in the SDB sample.

F.2 Calculation of Wave 4 longitudinal weights

As part of the BNLA methodology, once participants were recruited into the study at Wave 1, participants who could not be contacted or refused to participate at a particular wave could complete an interview at a later wave. In previous waves, survey weights were calculated for every possible combination of responses (Waves 1 and 3; Waves 2 and 3). This approach means that there is an increasing number of survey weights as each wave of data collection is completed. For example, Wave 4 with a separate weight for PA, SA and all participants and each combination of survey responses (Wave 1 and Wave 4; Wave 2 and Wave 4; Wave 3 and Wave 4) would require a set of 9 separate survey weights. From this wave, a shorter set of survey weights were calculated. This consisted of 6 survey weights calculated for Wave 4 covering two combinations of responses:

- Respondents to Waves 1, 2, 3 and 4 (longitudinal survey weight for respondents to all waves or 'balanced panel weight'. A separate weight for PAs, SAs and all participants are calculated).
- Respondents to Waves 1 and 4 (longitudinal survey weight for respondents from Wave 1 who completed a Wave 4 survey. A separate weight for PAs, SAs and all participants are calculated).

This general approach will be followed in future waves, with a set of 6 longitudinal survey weights supplied with each data release.

F.2.1 Factors predicting participation in all 4 waves

This section describes the calculation of the longitudinal survey weight for respondents to all waves, that is, the group of participants who were interviewed in Wave 1 and responded in each subsequent wave up to and including Wave 4. This group is counted 'as responses' and every other participant in Wave 1 who did not complete each subsequent wave as 'non-responses'. The analysis that follows identifies the characteristics of participants who responded in all 4 waves and these characteristics are used in the later survey weighting procedure to adjust for attrition across the 4 waves. Multivariate logistic models were estimated to identify the characteristics associated with survey completion. Separate models were estimated for PA, SA and all participants. Overall, 957 PAs and 592 SAs totalling 1,549 participants completed all four waves of data collection.

The analysis used to inform the calculation of longitudinal weights was based on the final Waves 1 to 4 datasets. As comprehensive data was collected on housing, English language proficiency, education and training, employment and income, health, self-sufficiency, community support and perceptions of life in Australia as part of the BNLA study data collection, a wide range of variables were considered in the following analysis. The following factors were tested to determine if they were significantly associated with survey completion across all 4 waves:

1. Visa sub-class, gender and age
2. Capital city
3. Household type

4. MU size
5. Marital status
6. Pre-migration education
7. SEIFA and remoteness index
8. Country of birth
9. English language proficiency
10. Employment
11. Mental health and posttraumatic stress
12. Whether waiting for family to migrate to Australia
13. Number of times moved home
14. Financial hardship and main source of income
15. Physical health

Statistical model results predicting participation in all 4 waves for each participant type are summarised in Table F.3. The model results show that PAs were more likely to participate across all 4 waves of data collection if they had higher levels of schooling, lower levels of understanding of spoken English and reported savings as their main source of income. PAs who arrived in Australia through the onshore migration pathway, were born in Iran, Africa or other countries were less likely to participate in all 4 waves of data collection.

A more limited set of factors predicted SAs participation in all 4 waves. Visa sub class and country of birth were the only two factors associated with SA survey completion. The model results for SA participation in all 4 waves mirrored the PA results, however SAs born in the Rest of Asia were significantly more likely to participate across all 4 waves, whereas PAs born in these countries were not significantly more likely to participate in the first 4 waves of interviews.

Table F.3: Results of multivariate logistic regression modelling survey participation in the first four waves, by participant type at Wave 1, odds ratio

BNLA Wave 1 characteristic	PA model	SA model	All participants
Pre-migration education			
Never attended school	1 (Base)		1 (Base)
6 or less years of schooling	1.525*		1.256
7-9 years of schooling	1.778*		1.573*
10-11 years of schooling	1.401		1.497*
12 or more years of schooling	1.813*		1.251
Trade or technical qualifications beyond school	2.151*		1.403
University Degree	1.722*		1.448
Don't know/prefer not to say	1.717		1.434
Age (years)	1.019***		1.013**
Visa Sub-Class			
200	1 (Base)	1 (Base)	1 (Base)
201	0.259	1.341	0.510
202	0.573	0.681	0.621
203			
204	0.784	1.161	0.950
866-non-UMA	0.589**	0.299*	0.508***
866-UMA	0.489*	1.135	0.590*
Understanding of spoken English			
Very well/well	1 (Base)		1 (Base)
Not well	1.205		1.156
Not at all	1.605*		1.383*
Don't know/prefer not to say	0.746		0.895
Country of birth			
Afghanistan	1 (Base)	1 (Base)	1 (Base)
Iran	0.515**	1.182	0.729
Iraq	1.362	2.500*	1.667***
Burma/Myanmar	0.886	1.333	0.951
Rest of mid-East	1.372	0.568	0.964
Rest of Asia	1.200	5.366***	1.800***
Africa	0.448*	0.276**	0.373***
Rest of world	0.356*	-	0.206***
Main income source			
Own wage or salary	1 (Base)		
Spouse/partner's/parent's income	2.101		
Financial support from others	1.296		
Government payments	1.001		
Savings	9.451*		
Other	2.003		
Don't know/prefer not to say	1.477		
Constant	**0.535**	**1.310**	**0.773**

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

F.2.2 Factors predicting participants from Wave 1 completing a Wave 4 interview

A second set of longitudinal weights were also calculated. These weights adjusted for attrition between Waves 1 and 4. Separate models were estimated to identify the factors associated with survey participation in Wave 4 by PAs and SAs. In total, 1,929 participants from Wave 1 also completed a Wave 4 interview. This comprised 1,186 PAs and 743 SAs. The same set of factors that were tested in the models predicting survey participation across all 4 waves described in the previous section were also tested in these models. The model results for PAs, SAs and all participants are shown in Table F.4.

The statistical modelling shows that PAs recruited in Wave 1 were more likely to participate in Wave 4 if they were born in Iraq and had arrived in Australia as part of Migrating Unit consisting of a couple family with children under 18 years old and other family members. Those born in Iran or Africa, arriving with an 866 non-UMA visa class and having family overseas were less likely to participate in a Wave 4 interview.

Similar to PA participation, SAs born in Africa were less likely to participate in Wave 4. Remoteness area and number of financial hardship events significantly predicted SA participation (but were not significantly associated with PA participation in Wave 4). SAs living in inner regional areas were more likely to complete a Wave 4 survey compared to SAs living in major cities. SAs reporting a higher number of financial hardship events were less likely to participate compared to those SAs reporting no financial hardship.

Table F.4: Results of multivariate logistic regression modelling predicting survey participation in Wave 4, by participant type at Wave 1, odds ratio

BNLA Wave 1 characteristic	PA model	SA model	All participants
Visa Sub-Class			
200	1 (Base)		1 (Base)
201	1.970		2.878
202	0.495		0.493*
203			
204	1.250		1.196
866-non-UMA	0.532**		0.489***
866-UMA	0.778		0.785
Remoteness area			
Major cities of Australia		1 (Base)	1 (Base)
Inner regional Australia		3.838**	2.186**
Outer regional Australia		0.195	0.768
Migrating unit structure			
Couple family with children under 18 (no other family members)	1 (Base)		1 (Base)
Couple family with children under 18 and other family members	3.806**		2.087**
Couple family only (no other family members)	1.261		1.140
Couple family and other family members but no children under 18	1.642		1.103
Single parent family with children under 18 (no other family members)	0.872		0.844
Single parent family with children under 18 and other family members	1.701		1.203
Other immediate family members	1.103		0.808
Other extended family members only			
Non-related persons			
Single person	0.788		0.749
Country of birth			
Afghanistan	1 (Base)	1 (Base)	1 (Base)
Iran	0.653*	0.790	0.684*
Iraq	1.563*	1.657	1.547**
Burma/Myanmar	1.355	0.520	1.009
Rest of mid-East	0.785	0.361	0.577
Rest of Asia	0.798	2.002	0.918
Africa	0.468*	0.138**	0.234***
Rest of world	0.398	0.203	0.297**
Has family overseas waiting to come to Australia			
Yes	1 (Base)		1 (Base)
No	0.706*		0.748*
Don't know / prefer not to say	0.726		0.887
Number of financial hardship events			
None		1 (Base)	
1		0.949	
2		0.813	
3		2.678	
4		0.707	
5		0.173*	
6			
Don't know / prefer not to say		1.154	
Constant	**4.910*** **	**4.803*** **	

Notes: * p < 0.05, ** p < 0.01, *** p < 0.001.

F.2.3 Calculation of balanced panel and Waves 1 and 4 longitudinal survey weights

As before, the appropriateness of generating a single longitudinal weight across both PA and SA participants was further investigated within a variance-component modelling framework. Model results indicated that it was appropriate to calculate both a single balanced panel longitudinal weight and single Wave 1/4 longitudinal survey weight (less than 5% of the variance in survey completion was explained at the individual level in both cases).

The raking procedure described earlier was used to adjust for the differential pattern in non-response of the recruited sample who completed the first four waves of interviews compared to those who missed at least one interview wave. The same analysis was used to adjust for differential pattern in the recruited sample who completed an interview in Wave 4. Longitudinal survey weights were calculated using the raking procedure. A description of these survey weights and the variables used in the survey weighting procedure for each of these variables can be found in the next section.

F.3 Survey weights available in the Wave 4 dataset

In total, nine survey weights are available in the Wave 4 BNLA dataset. Three 'population' and six 'longitudinal' survey weights are provided in the dataset. These survey weights and their variable names are described further below:

Population weights

1. ***dpawgt*** - Wave 4 Principal Applicant Survey weight. The following variables were used to develop this survey weight variable: visa sub-class, capital city, family size and country of birth. The calculated survey weights ranged in value from 0.31 to 4.83. After an examination of the distribution and factors associated with the larger value survey weights, the calculated survey weights were capped at a maximum value of 2.5.

   It is recommended that this survey weight be used for cross-sectional analysis involving questions only asked of Principal Applicants in Wave 4.

2. ***dsawgt*** - Wave 4 Secondary Applicant survey weight. The visa sub-class, major city, country of birth and migrating unit structure variables were used to develop survey weights for this variable. SA survey weights ranged from 0.31 to 5.60. As with the PA weight, the SA weight was constrained to take a maximum value of 2.5.

   It is recommended that this survey weight be used for cross-sectional analysis involving questions asked only of Secondary Applicants in Wave 4.

3. ***dwgt*** – Wave 4 all participant survey weight. Visa sub-class, capital city, age and country of birth information were used in the survey weighting for this variable. This survey weight ranged from 0.45 to 3.10 and was also capped at a maximum value of 2.5.

It is recommended that this survey weight be used when analysing survey questions asked of both PA and SA respondents in Wave 4.

Longitudinal weights

1. ***a_dpawgt*** – Waves 1, 2, 3 and 4 balanced panel longitudinal weight for Principal Applicants. The following variables were used to develop this survey weight variable: visa sub-class, country of birth, pre-migration education, age, English language proficiency and main source of income. The calculated survey weights ranged in value from 0.69 to 1.83.
2. ***a_dsawgt*** – Waves 1, 2, 3 and 4 balanced panel longitudinal survey weight for Secondary Applicants. The visa sub-class and country of birth variables were used to develop this survey weight. Wave 1-4 SA longitudinal survey weights ranged from 0.84 to 1.82.
3. ***a_dwgt*** – Waves 1, 2, 3 and 4 balanced panel longitudinal survey weight for all participants. Visa sub-class, country of birth, age, pre-migration education, and English language proficiency information were used in the survey weighting for this variable. This survey weight ranged from 0.71 to 1.84.
4. ***adpawgt*** – Wave 1/4 Principal Applicant longitudinal survey weight. The following variables were used to develop this survey weight variable: visa sub-class, country of birth, migrating unit structure and family overseas waiting to come to Australia. The calculated survey weights ranged in value from 0.76 to 1.44.
5. ***adsawgt*** – Wave 1/4 Secondary Applicant longitudinal survey weight. The country of birth, remoteness area and number of financial hardship events, variables were used to develop this survey weight variable. Wave 1/4 SA longitudinal survey weights ranged from 0.88 to 1.62.
6. ***adwgt*** – Wave 1/4 all participant longitudinal survey weight. Visa sub-class, country of birth, migrating unit structure, family overseas and remoteness area information were used to derive this survey weight variable. This survey weight ranged from 0.75 to 1.47.

Appendix G: Wave 5 survey weighting

This section describes the process used to calculate survey weights for the Wave 5 sample of Building a New Life in Australia. The same processes was followed to calculate Wave 5 survey weights as for previous waves (See Appendix C, D E, and F for further information on the BNLA survey weighting approach).

Two types of survey weights were calculated and are available on the Wave 5 dataset:

- A 'population' weight – that adjusts BNLA Wave 5 estimates to population totals (in this case, the population of humanitarian migrants who have been granted a permanent visa or arrived in Australia between May and December 2013).[72]
- A 'longitudinal' weight that adjusts for attrition between the first and fifth wave, and a balanced panel survey weight for respondents who completed all of the first five waves of data collection. For this wave, a longitudinal weight was also calculated for those who participated in Waves 1, 3 and 5 to reflect those waves where interviews were administered in the home.

G.1 Calculation of Wave 5 population weights

Following the same processes as in previous waves, analysis of non-response was undertaken to identify the characteristics of those who completed an interview compared with all humanitarian migrants on the SDB that were eligible to participate in the study at the time of study recruitment.[73] The following section describes the results of multivariate logistic models. These models were estimated to identify the factors associated with PA participation, SA participation, and survey completion by all participants in Wave 5.

Factors associated with participation in Wave 5

Results of statistical models predicting Wave 5 participation for each participant type are summarised in Table G.1.

[72] Offshore visa holders had to have **arrived** in Australia holding a permanent visa, while onshore visa holders had to have **received** their protection visa in this time period to be eligible for participation in BNLA.
[73] n=7,362 individuals in 4,035 migrating units were eligible for selection. For more information about the sampling and sampling frame refer to the Sections 2.1 and 2.2 of this Data Users Guide.

Table G.1: Results of multivariate logistic regression predicting survey participation in Wave 5, by participant type, odds ratios

SDB characteristic	PA model	SA model	All participants
Gender			
Male	1.104	0.817	1.070
Female	1 (Base)	1 (Base)	1 (Base)
Site			
Site 1#	0.533*	0.544	0.547*
Site 2#	0.258***	0.232***	0.261***
Site 3	0.574	0.424	0.544
Site 4#	0.372*	0.906	0.428
Site 5#	0.770	0.394	0.580
Site 6#	0.451**	0.222***	0.342***
Site 7	0.382*	-	0.253**
Site 8#	0.310***	0.192**	0.256***
Site 9#	0.238***	0.219***	0.232***
Site 10	0.526	0.088***	0.240***
Site 11	1 (Base)	1 (Base)	1 (Base)
Visa Sub-Class			
200	1 (Base)	1 (Base)	1 (Base)
201	1.528	1.074	1.402
202	0.887	0.924	0.935
203	-	-	-
204	1.184	1.580	1.461**
866-non-UMA	0.362***	0.242***	0.294***
866-UMA	1.185	0.831	0.944
Family size	1.209***		
Country of birth			
Afghanistan	1 (Base)	1 (Base)	1 (Base)
Iran	1.086	1.085	1.071
Iraq	1.960***	1.122	1.555***
Burma/Myanmar	0.309***	0.145***	0.243***
Rest of mid-East	0.943	0.430*	0.741
Rest of Asia	0.746	0.584*	0.707*
Africa	0.387***	0.201***	0.315***
Rest of world	0.545	-	0.317
Marital Status			
Divorced	1 (Base)		1 (Base)
Engaged	2.455		1.965
Defacto	2.427		2.567
Married	1.290		1.457
Never married	1.252		0.838
Separated	1.564		1.611
Widowed	1.598		1.327
Household type			
Male PA		1 (Base)	
Male PA / married female SA		1.914**	
Male PA / unmarried male or female SA		1.043	
Male PA / Child 15-17 years		1.128	
Constant	0.594	1.488	

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Gender included as a control variable.
# indicates site is a capital city. Location of sites have not been named to protect participant confidentiality.

The model results show that both PA and SA participants shared some factors that predict survey completion in Wave 5. For both the PA and SA sample, survey participation was significantly associated with region, visa subclass and their country of birth. PAs who were born in Iraq were more likely to participate compared to those born in Afghanistan, however there was no significant association with survey participation for SAs who were born in Iraq. Both PA and SA participants born in Burma/Myanmar and Africa were less likely to participate also in Wave 5.

Family size was significantly associated with PA participation in Wave 5. PAs with larger families were more likely to complete a Wave 5 survey, compared to those with smaller families. Household type was significantly associated with SA participation but was not found to be significant for PA participants.

G.2 Calculation of Wave 5 cross-sectional survey weights

As in previous waves, further analysis was undertaken to identify the appropriateness of calculating a single survey weight across PA and SA respondents using variance component modelling. The rationale for this approach and further description of the variance component modelling is provided in Appendices C through F. As in previous waves, the variance component models showed that low levels of the variance in survey completion was found at the individual level (around 10%). As most of the variance in survey completion could be explained at the migrating unit level, a single Wave 5 population weight for all participants was also calculated following the raking procedure described in previous Appendices. The variables found to be significantly associated with survey participation described in the previous section were used in the raking procedure to generate the PA, SA and overall participant population weight. These survey weights adjust the survey estimates to achieve population totals, in this case the population of humanitarian migrants being granted a permanent visa between May and December 2013.[74]

Outcomes of Wave 5 survey weighting

The effects of survey weighting were further explored by comparing unweighted and weighted estimates of the Wave 5 BNLA data with Settlement Database (SDB) variables not used in the weighting procedure (gender, and age group). As Table G.2 shows, the estimate for age was more accurate when survey weights were used, namely the survey weights brought the BNLA population estimate closer to the SDB total population proportions.

However, as in previous waves, the estimate for gender was slightly less accurate when the survey weighted estimate was compared to the unweighted estimate.

[74] Offshore visa holders had to have **arrived** in Australia holding a permanent visa, while onshore visa holders had to have **received** their protection visa in this time period**.**

Table G.2: Effect of survey weighting on estimates of variables in the Wave 5 BNLA data and SDB population sample

	Principal Applicants		
	BNLA Wave 5 unweighted	**BNLA Wave 5 weighted**	**SDB Sample**
Gender			
Male	69.7%	70.8%	69.7%
Female	30.3%	29.2%	30.3%
Age Group[(a)]			
18-30 years	26.5%	29.2%	32.8%
31-55 years	60.8%	58.8%	55.1%
56 years or over	12.7%	12.1%	12.1%
Total	**1,186**	**1,186**	**4,035**

Notes: (a) Information on age group was missing for 0.5% of records in the SDB sample.

G.3 Calculation of Wave 5 longitudinal weights

As described in Appendix F, once recruited into the study at Wave 1, participants who refused to participate or could not be contacted at a particular wave could re-join the study at a later wave. Due to the study methodology, this means there is an ever increasing number of survey weights that potentially could be calculated as each data collection wave is completed. For example, in Wave 5 if a separate weight for each participant type (PA, SA and all participants) and each combination of survey weights (Wave 1 and Wave 5, Wave 1,2 and 5, Wave 1,2,3 and 5, Wave 1,2,4,5, Wave 2,3,4,5 etc), were to be calculated – this would require the generation of 48 separate survey weights (16 combinations of survey weights x 3 applicant types). In this wave, a shorter set of survey weights were calculated. This involved the calculation of 9 survey weights for Wave 5 covering three combination of responses:

- Respondents to Waves 1, 2, 3, 4 and 5 (longitudinal survey weight for respondents to all waves or 'balanced panel weight'. A separate weight for PAs, SAs and all participants are calculated).
- Respondents to Waves 1 and 5 (longitudinal survey weight for respondents from Wave 1 who completed a Wave 5 survey. A separate weight for PAs, SAs and all participants are calculated).
- Respondents to Waves 1, 3 and 5 (longitudinal survey weight for respondents from Wave 1 who completed a Wave 3 and Wave 5 survey. A separate weight for PAs, SAs and all participants are calculated.

Factors predicting participation in all 5 waves

This section describes the calculation of the longitudinal survey weight for respondents to all waves, that is, the group of participants who were interviewed in Wave 1 and responded in each subsequent wave up to and including Wave 5. This group is counted 'as responses' and every other participant in Wave 1 who did not complete each subsequent wave as 'non-responses'. The analysis that follows identifies the characteristics of participants who responded in all 5 waves and these characteristics are used in the later survey weighting procedure to adjust for attrition across the 5 waves. Multivariate logistic models were

estimated to identify the characteristics associated with survey completion. Separate models were estimated for PA, SA and all participants. Overall, 1,447 participants completed all 5 waves. This comprised 886 PAs and 561 SAs.

The analysis used to inform the calculation of longitudinal weights was based on the Waves 1 through 5 datasets. As in previous waves, the same factors were tested to determine if they were significantly associated with survey completion across all 5 waves:

- Visa sub-class, gender and age
- Capital city
- Household type
- MU size
- Marital status
- Pre-migration education
- SEIFA and remoteness index
- Country of birth
- English language proficiency
- Employment
- Mental health and posttraumatic stress
- Whether waiting for family to migrate to Australia
- Housing tenure
- Number of times moved home
- Financial hardship and main source of income
- Physical health

The model results predicting participation in all 5 waves for each participant type are summarised in Table G.3. The model results show that PAs were more likely to participate across all 5 waves of data collection if they had higher levels of schooling, born in Iraq and were older. PAs who arrived in Australia through the onshore migration pathway, or were born in Africa were less likely to participate in all 5 waves.

In terms of SA participation, visa sub class and country of birth were significantly associated with participation in all 5 waves. The model results for SA participation were very similar to the factors associated with PA participation, however SA participants born in the rest of Asia were significantly associated with participation in all 5 waves, while this result did not hold for PA participants.

Table G.3: Results of multivariate logistic regression modelling survey participation in all 5 waves, by participant type at Wave 1, odds ratios

BNLA Wave 1 characteristic	PA model	SA model	All participants
Gender			
Male	0.874	0.868	0.940
Female	1 (Base)		1 (Base)
Pre-migration education			
Never attended school	1 (Base)		1 (Base)
6 or less years of schooling	1.468		1.235
7-9 years of schooling	1.465		1.438*
10-11 years of schooling	1.156		1.259
12 or more years of schooling	1.363		1.006
Trade or technical qualifications beyond school	1.396		1.118
University Degree	1.127		1.107
Don't know/prefer not to say	1.510		1.089
Age (years)	1.021***		1.015***
Visa Sub-Class			
200	1 (Base)	1 (Base)	1 (Base)
201	0.249	2.143	0.433
202	0.702	0.896	0.729
203	-	-	-
204	0.791	1.356	1.042
866-non-UMA	0.472	1.608	0.568*
866-UMA	0.638	0.319	0.553***
Country of birth			
Afghanistan	1 (Base)	1 (Base)	1 (Base)
Iran	0.661	1.162	0.846
Iraq	1.890***	2.718***	2.091***
Burma/Myanmar	0.953	1.172	0.949
Rest of mid-East	1.821	0.498	1.195
Rest of Asia	1.288	4.393**	1.763**
Africa	0.351**	0.340*	0.337***
Rest of world	0.501	-	0.270**
Constant	**0.543****	**1.077**	**0.679**

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Gender included as a control variable.

G.3.2 Factors predicting participants from Wave 1 completing a Wave 5 interview

A set of survey weights were also calculated to adjust for attrition between Waves 1 and 5. Separate statistical models were estimated to identify the factors associated with survey participation in Wave 5 for PA, SA and all participants. Overall, 1,881 participants from Wave 1 also completed an interview in Wave 5. This number was made up of 1,144 PAs and 737 SAs. Model results showing the factors associated with survey participation in Wave 5 by applicant type are shown in Table G.4.

The factors associated with PA participation in Wave 5 include: country of birth (being born in Iraq), higher levels of pre-migration education, having arrived in Australia through an offshore migration pathway and arriving in Australia as part of a Migrating Unit consisting of a single parent family with children under 18 years old and other family members.

SA participation in Wave 5 was also significantly associated with country of birth and visa sub-class. There were some contrasting results with the PA model results, pre-migration education levels were not significantly associated with SA participation, however, understanding of spoken English was a significant predictor of participation for SAs. Those with lower levels of English understanding were more likely to complete a Wave 5 interview.

Table G.4: Results of multivariate logistic regression models predicting survey participation in Wave 5, by participant type at Wave 1, odds ratios

BNLA Wave 1 characteristic	PA model	SA model	All participants
Gender			
Male	1.126	0.774	1.010
Female	1 (Base)	1 (Base)	1 (Base)
Visa Sub-Class			
200	1 (Base)		1 (Base)
201	1.827	-	3.150
202	0.986	0.793	0.890
203	-	-	-
204	1.341	2.995*	2.052**
866-non-UMA	0.422**	0.452	0.474*
866-UMA	0.771	0.413*	0.772
Pre-migration education			
Never attended school	1 (Base)		
6 or less years if schooling	1.565*		
7-9 years of schooling	1.407		
10-11 years of schooling	1.053		
12 or more years of schooling	1.180		
Trade or technical qualifications beyond school	0.904		
University Degree	0.999		
Don't know/prefer not to say	1.672		
Understanding of spoken English			
Very well		1 (Base)	1 (Base)
Well		2.778*	1.353
Not well		2.381*	1.522
Not at all		2.752*	1.644*
Don't know/prefer not to say		2.491	1.962
Migrating unit structure			
Couple family with children under 18 (no other family members)	1 (Base)		1 (Base)
Couple family with children under 18 and other family members	1.580		1.409
Couple family only (no other family members)	1.132		1.010
Couple family and other family members but no children under 18	1.003		0.699
Single parent family with children under 18 (no other family members)	0.961		0.624
Single parent family with children under 18 and other family members	3.303*		1.287
Other immediate family members	0.883_		0.476*
Other extended family members only	_		-
Non-related persons	0.624*		-
Single person			0.574*
Country of birth			
Afghanistan	1 (Base)	1 (Base)	1 (Base)
Iran	1.089	1.182	0.993
Iraq	2.530***	2.987**	2.551***
Burma/Myanmar	1.088	0.592	0.838
Rest of mid-East	21.73	0.806	1.182
Rest of Asia	1.255	3.333*	1.558*
Africa	0.622	1.101	0.724
Rest of world	0.478	0.090*	0.248**
Constant	**2.283***	**1.294**	**2.283****

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Gender included as a control variable.

Factors predicting participants from Wave 1 completing a Wave 3 and Wave 5 interview

For this wave, we also undertook analysis of the factors associated with survey participation in Waves 1, 3 and 5. This analysis can help to identify the characteristics of participants who participated in all of the home visit interviews. Overall, 1,662 participants (1,000 PAs; 662 SAs) completed all three home visit interviews.

The results in Table G.5 model the probability of completing a Wave 1, 3 and 5 survey relative to the group who did not complete at least one of the home visit interviews. The model results show that for all participants, visa sub-class, age, pre-migration education, migrating unit structure and country of birth were all significantly associated with completion of Waves 1, 3 and 5 surveys. Participants who were older, had higher levels of pre-migration education, and were born in Iraq or the rest of Asia were all factors associated with home visit survey completion. Conversely arriving in Australia as part of a migrating unit as a single person or with other immediate family members or via an onshore migration visa, as was being born in Africa or the rest of the world were factors that were found to lower the likelihood of participation in Waves 1, 3 and 5. The same set of factors predicted PA participation in all three home visit waves. Analysis of SA participation showed that compared with PA and all participants, migrating unit structure and age was not significantly associated with SA participation in the three home visit surveys, however, SAs who were partnered at Wave 1 were more likely to participate in Waves 1, 3 and 5 compared to those who were unpartnered.

Table G.5: Results of multivariate logistic regression models predicting survey participation in Wave 1 and Wave 3 and Wave 5, by participant type at Wave 1, odds ratios

BNLA Wave 1 characteristic	PA model	SA model	All participants
Gender			
Male	1.015	0.834	0.871
Female	1 (Base)		1 (Base)
Visa Sub-Class			
200	1 (Base)	1 (Base)	1 (Base)
201	0.701	1.809	0.942
202	0.557	0.675	0.640
203	-	-	-
204	1.012	2.125*	1.326
866-non-UMA	0.502*	0.799	0.575*
866-UMA	0.762	0.275	0.683*
Age (years)	1.023*		1.019*
Partnered			
No		1 (Base)	
Yes		1.843**	
Don't know/prefer not to say		1.950*	
Pre-migration education			
Never attended school	1 (Base)	1 (Base)	1 (Base)
6 or less years of schooling	1.934**	1.045	1.667**
7-9 years of schooling	1.691*	0.774	1.422
10-11 years of schooling	1.341	0.797	1.211
12 or more years of schooling	1.594*	0.504*	1.066
Trade or technical qualifications beyond school	1.109	0.435	0.866
University Degree	1.202	1.131	1.123
Don't know/prefer not to say	1.754	0.504	1.156
Migrating unit structure			
Couple family with children under 18 (no other family members)	1 (Base)		1 (Base)
Couple family with children under 18 and other family members	1.008		0.871
Couple family only (no other family members)	0.781		0.710
Couple family and other family members but no children under 18	0.837		0.747
Single parent family with children under 18 (no other family members)	1.120		0.715
Single parent family with children under 18 and other family members	3.967**		1.599
Other immediate family members	0.857		0.555*
Other extended family members only	-	-	-
Non-related persons	-		-
Single person	0.670*		0.536***
Country of birth			
Afghanistan	1 (Base)	1 (Base)	1 (Base)
Iran	0.892	1.253	0.960
Iraq	1.901**	2.187*	1.856***
Burma/Myanmar	0.859	0.807	0.767
Rest of mid-East	1.499	0.403	0.811
Rest of Asia	1.156	2.809*	1.460*
Africa	0.421**	0.597	0.475**
Rest of world	0.361*	-	0.158***
Constant	**0.630**	**1.898**	**1.203**

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Gender included as a control variable.

Calculation of longitudinal survey weights

The raking procedure described earlier was used to adjust for the differential pattern in non-response of the recruited sample who completed all five waves (the 'balanced panel weight). The same raking analysis was also used to calculate longitudinal survey weights for those who completed a Wave 1 and 5 interview, and separately for those who completed an interview in Waves 1, 3 and 5. A description of these survey weights and the variables used in the survey weighting procedure for each of these variables can be found in the next section.

G.4 Survey weights available in the Wave 5 dataset

In total, twelve survey weights are available in the Wave 5 BNLA dataset. Three 'population' and nine 'longitudinal' survey weights are provided in the dataset. These survey weights and their variable names are described further below:

Population weights

1. ***epawgt*** – Wave 5 PA Survey weight. The following variables were used to develop this survey weight variable: visa sub-class, region (capital city), family size, marital status and country of birth. The calculated survey weights ranged in value from 0.31 to 5.73. After an examination of the distribution and factors associated with the larger value survey weights, the calculated survey weights were capped at a maximum value of 2.5.

   It is recommended that this survey weight be used for cross-sectional analysis involving questions only asked of PAs in Wave 5.

2. ***esawgt*** – Wave 5 SA survey weight. The visa sub-class, region (capital city), country of birth and household structure variables were used to develop survey weights for this variable. SA survey weights ranged from 0.23 to 8.37. As with the PA weight, the SA weight was constrained to take a maximum value of 2.5.

   It is recommended that this survey weight be used for cross-sectional analysis involving questions asked only of SAs in Wave 5.

3. ***ewgt*** – Wave 5 all participant survey weight. Visa sub-class, region (capital city), marital status and country of birth information were used in the survey weighting for this variable. This survey weight ranged from 0.38 to 5.80 and was also capped at a maximum value of 2.5.

   It is recommended that this survey weight be used when analysing survey questions asked of both PA and SA respondents in Wave 5.

Longitudinal weights

1. ***a_epawgt*** – Waves 1, 2, 3, 4 and 5 balanced panel longitudinal weight for Principal Applicants. The following variables were used to develop this survey weight variable: visa sub-class, country of birth, pre-migration education, and age. The calculated survey weights ranged in value from 0.67 to 2.02.

2. ***a_esawgt*** – Waves 1, 2, 3, 4 and 5 balanced panel longitudinal survey weight for Secondary Applicants. The visa sub-class and country of birth variables were used to

develop this survey weight. Wave 1-5 SA longitudinal survey weights ranged from 0.83 to 1.86.

3. ***a_ewgt*** – Waves 1, 2, 3, 4 and 5 balanced panel longitudinal survey weight for all participants. Visa sub-class, country of birth, age, and pre-migration education, were used in the survey weighting for this variable. This survey weight ranged from 0.72 to 2.92.
4. ***aepawgt*** – Wave 1/5 Principal Applicant longitudinal survey weight. The following variables were used to develop this survey weight variable: visa sub-class, country of birth, migrating unit structure and pre-migration education. The calculated survey weights ranged in value from 0.65 to 1.33.
5. ***aesawgt*** – Wave 1/5 Secondary Applicant longitudinal survey weight. The visa sub-class, country of birth, and understanding of spoken English variables were used to develop this survey weight variable. Wave 1/5 SA longitudinal survey weights ranged from 0.90 to 1.50.
6. ***aewgt*** – Wave 1/5 all participant longitudinal survey weight. Visa sub-class, country of birth, migrating unit structure, and understanding of spoken English were used to derive this survey weight variable. This survey weight ranged from 0.74 to 1.68.
7. ***acepawgt*** – Wave 1/3/5 Principal Applicant longitudinal survey weight. The visa sub-class, age, migrating unit, country of birth and pre-migration education variables were used to generate this weight. This survey weight ranged from 0.59 to 2.28.
8. ***acesawgt*** – Wave 1/3/5 Secondary Applicant longitudinal survey weight. Visa sub-class, partnered, pre-migration education and country of birth information were used to derive this weight. This survey weight ranged from 0.64 to 1.90.
9. ***acewgt*** – Wave 1/3/5 all participant longitudinal survey weight. Visa sub-class, age, migrating unit, country of birth and pre-migration education variables were used to derive this weight. This survey weight ranged from 0.62 to 2.50.