

Paper delivered at the 10<sup>th</sup> Biennial Conference of the  
*Australian Population Association*  
POPULATION AND GLOBALISATION:  
AUSTRALIA IN THE 21<sup>ST</sup> CENTURY  
Melbourne 28<sup>th</sup> November to 1<sup>st</sup> December 2000  
Melbourne Australia

Yin Paradies, Sarah Huppertz, Jim Warnsey and Tony  
Barnes, Population estimates for remote area Indigenous  
communities

*Yin Paradies, Sarah Huppertz, Jim Warnsey and Tony Barnes, Population estimates for remote area Indigenous communities*

## **Introduction**

I would like to begin by acknowledging the the traditional owners of the land on which we are gathered today.

## **Slide 1: Title Slide**

I am a project development officer with the Cooperative Research Centre for Aboriginal and Tropical Health in Darwin and I will be talking today about a pilot project undertaken by the CRCATH on estimating populations in remote Indigenous communities.

## **Project Background**

Currently counts from the ABS Census of Population and Housing are the only comprehensive systematically collected remote area community population figures that are collected using a formal and well-established methodology. This data is collected once every five years and community level counts are not released by the ABS as population estimates because it is well recognised that counts may not provide a complete enumeration.

As you are well aware annual population figures are key determinants of resource allocation and are essential to the accurate monitoring of health outcomes of which the CRCATH has a particular interest.

Population figures at intervals more frequent than every 5 years are sometimes required for individual communities to monitor population movements and to determine the need for health services and infrastructure.

At the present time there is no widely accepted satisfactory method available for deriving reliable population estimates for remote Indigenous communities by either adjusting Census counts or by other means.

In response to this situation, the Cooperative Research Centre for Aboriginal and Tropical Health (CRCATH) decided to conduct preliminary research into a methodology which, if proven reliable and feasible, would provide a method for deriving population estimates in remote communities.

The project involves field testing of an experimental sampling procedure in parallel with extensive consultations with stakeholders such Indigenous communities and organisations as well as Commonwealth and Territory government agencies. A consultative approach to the research was adopted to enable all interested parties to become stakeholders in the project by contributing their ideas at the design stage and assisting in the identification of the resources necessary to implement any latter stages of the project.

## **Slide 2: Situation**

The situation and question being explored can be shown in this diagram.

## **Slide 3: Project Methodology**

A theoretical method for deriving the size of community populations has been shown on this slide. I will go through the maths later but the method can be summarised in simple terms as follows:

## **Slide 4: LIST & AREA FRAMES**

A list of the names of potential community residents is prepared using the best information available. The list must contain sufficient information about people to ensure that they can be identified with certainty and that there are no duplications. The list need not be complete and may contain names in excess of the people likely to be found in the community at a particular future point in time.

Using SLAP (Serviced Land Availability Plan) maps, aerial photography and local community knowledge of all the houses and improvised dwellings which could have potential residents. These dwellings were identified and organised into a frame that could then be sample from at a future date.

## **Slide 5: SAMPLING PROCEDURES**

At a specific time for which a population estimate is required, a random sample of names is taken from the list and an attempt is made to match them against people actually found to be present in the community. At the same time, a number of people in the community are selected by cluster sampling of people in dwellings.

## **Slide 6: MATCHING**

These people are then identified and matched, if possible, against the names on the list.

## **Slide 7-9: THEORY**

An unbiased point estimate (and approximate confidence interval) of the number of people present in the community at the time of sampling may be calculated from the number of people on the list, the number of people in each sample and the number of matches for each sample.

## **Slide 10: PRECISION OF ESTIMATE**

We also derived a relative standard error for the estimate as shown here.

### **Slide 11: PRACTICALITIES**

There were a number of practical issues that we wished to explore as part of the pilot project including those listed on this slide.

I don't have time to discuss these issues in any depth but essentially we found very positive results to all of these issues from the pilot.

Community leaders were very supportive of the project.

It was possible to construct the list by talking to heads of families within the community and constructing family trees and then converting to a list.

Random and cluster sampling and matching were all possible.

There were also a number of other technical and practical issues that arose during the pilot some of which have already been resolved. I will briefly describe some further issues which need exploring shortly.

### **Slide 12: ISSUES TO EXPLORE**

These are further issues that need to be explored before and during further uses of this methodology in consultation with interested stakeholders.

### **Slide 13: FUTURE OPTIONS**

And, in conclusion, we hope to explore a number of options for future uses of this methodology in association with interested stakeholders.

The first two options would seem to be the most beneficial ways of applying this methodology. The first being the Census adjustment discussed earlier and the second involves producing sequential population estimates for high mobility populations such as town camps and Indigenous communities and outstations by repeating sampling at different points in time using a common list that would be constructed at the outset.



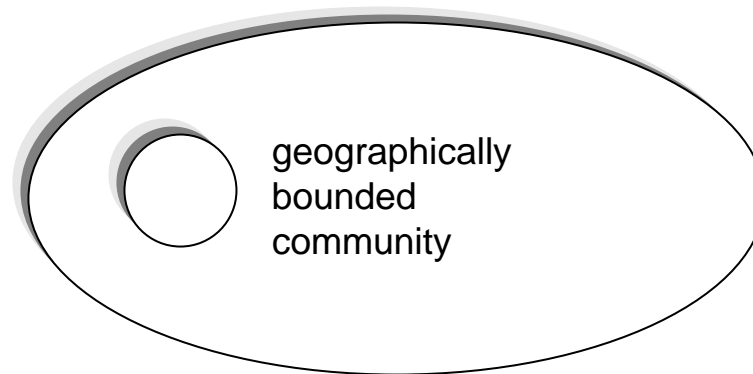
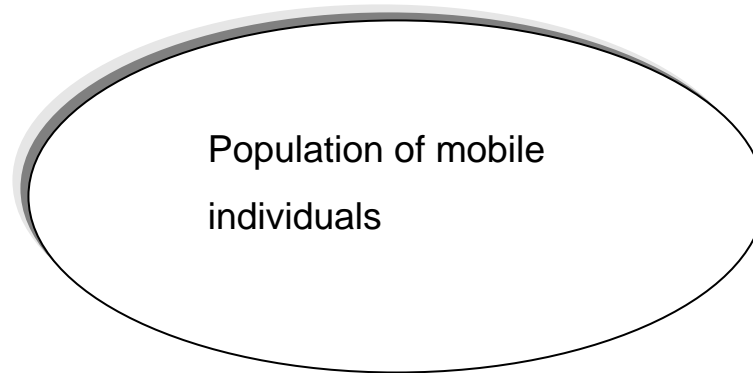


# **INVESTIGATION OF A METHODOLOGY FOR ESTIMATING COMMUNITY POPULATIONS**

**Yin Paradies, Sarah Huppatz, Jim Wansey, Tony  
Barnes**

**November 2000**

# Situation



## Question

How many people in the population are currently in the Community?

# REQUIREMENTS OF ESTIMATION PROCEDURE

## A. TWO SAMPLING FRAMES

1. List frame of individuals in the population

- need not be complete

2. Area frame of individuals (households) currently in community

- must be complete

## B. TWO SAMPLING PROCEDURES

3. A random sample from the list frame

4. To select a random-based sample from the area frame

## C. TWO SEARCH AND MATCH PROCEDURES

5. Match individuals in the list sample with individuals in the community.

6. To match individuals in the area sample with individuals on the list frame

# DEVELOPMENT OF LIST & AREA FRAMES

- **LIST FRAME**
- **Key Information:**
  - knowledge held by community members about family/clan relationships
- **Capture of information:**
  - through engagement of community elders to develop genealogy of each family within the community.
- **Convert genealogy to list:**
  - included identifying information such as name, age, sex, and place of residence including house number where possible.
- **AREA FRAME**
- Using available maps and community knowledge construct an area frame of all residential households (including improvised dwellings and camps) in the community.

# SAMPLING

- **List sample:**

- A simple random sample of people.

- **Area Sample:**

- Cluster sampling by firstly taking a simple random sample of households and then enumerating all individuals in the household (along with necessary identifying information).

# MATCHING

- **List Sample to Community:**
  - enquire from senior family members about presence or absence of sampled individuals.
- **Area Sample to List:**
  - straight forward list matching.

# THEORY

Let  $P$  denote the community population which we are attempting to estimate and  $L$  denote the number of people in the list frame.

$P$  = the No. of people on the list who are present in the community  
+ the No. of people in the community who are not on the list

An estimate of the No. of people on the list who are currently in the community is:

$$L * (SA/SL)$$

$SL$  = No. of people sampled from the list frame

$SA$  = No. of people from the list sample who were present in the community

So that  $(SA/SL)$  is the proportion of people from the list sample who were present in the community

An estimate of the No. of people in the community who are not on the list is:

$$P * 1-(SC/SP)$$

Where,

SP = number of people in the household sample

SC = number of people from the household sample who were also on the list.

So that  $(SC/SP)$  is the proportion of people from the household sample who were found on the list.

And  $(1- SC/SP)$  is the proportion of people from the area sample who were not found on the list.

So an estimate of the population P is:

$$P' = L * (SA/SL) + P * (1 - SC/SP)$$

$$P' - P * (1 - SC/SP) = L * (SA/SL)$$

$$P' (1 - (1 - SC/SP)) = L (SA/SL)$$

$$P' (SC/SP) = L (SA/SL)$$

$$P' = L (SA/SL) / (SC/SP)$$

# PRECISION OF POPULATION ESTIMATE

Population estimate  $P' = L(SA/SL)/(SC/SP)$

Relative Standard Error of  $P'$  is approximately

$$RSE(P') = (1/SA - 1/SL + \Delta/SC - \Delta/SP)^{1/2}$$

Given that the list and area sampling were conducted by SRS and cluster sampling respectively, and where  $\Delta$  is the design effect of cluster sampling.

Note that the  $RSE(P')$  does not depend on list size.

# **PRACTICALITIES**

- **Will community leaders and organisations support the project?**
- **Is it possible to compile the list and if so, how?**
- **Is it possible to conduct random sampling (or equivalent) and matching with sufficient rigour?**
- **What resources are required to compile the list and to undertake sampling and are these forthcoming?**
- **What other practical and technical issues may arise as the method is applied?**

## **OUTCOME OF STAGE 1**

**Method has been found to be effective and potentially useable with support needed from external agencies to explore further issues and future options.**

# ISSUES TO EXPLORE:

- **What is the best approach to recording data in larger communities with a view to utilising genealogical/ethnographic/demographic techniques?**
- **Is this dual frame sampling and if so does this affect the methodology?**
- **Is it appropriate and if so what is the best way to tailor the list before each point in time estimate to produce higher match rates during sampling?**
- **How often and in what manner should lists be updated to account for births, deaths, and migration and should administrative sources be used?**
- **How do we handle community boundary and geographic definitional issues?**
- **Can this method provide accurate sequential (point-in-time) estimates?**
- **Does this method result in excessive respondent burden?**
- **Can this method be used to adjust remote area Census counts?**

# **FUTURE OPTIONS**

**Option 1: Census PES approach**

**Option 2: Conduct sequential population estimates**

**Option 3: Larger version of the pilot in 6+ communities**

**Option 4: Assist communities to develop their own lists**

**Option 5: List development in its own right**