Problem gambling subgroups in the Northern Territory

The multivariate risk factors for problem gambling (as defined by the Problem Gambling Severity Index [PGSI]) were documented in the Northern Territory Gambling Prevalence Survey 2005 (Young et al. 2006). These risk-factors included living in a group household, having lower educational outcomes, and regular poker-machine play. While this research was valuable in determining risk factors aggregated across the population, it did not account for the interaction between these variables in defining particular subgroups of problem gamblers, nor did it include spatial or behavioural variables. Therefore, based on more recent survey data, we set about the identification of these subgroups (for the full paper see Markham et al 2012).

Methods
Between April and August 2010 we conducted a mail survey of all NT households in the Geocoded National Address File (G-NAF: PSMA Australia, 2010) to which Australia Post would deliver unaddressed mail (n = 46,263). We also hand-delivered questionnaires to an additional 2,300 households in the peri-urban fringes of Darwin and Alice Springs. Any household member aged eighteen or older was eligible to respond. The Human Research Ethics Committee of Charles Darwin University granted approval to conduct the study (protocol no. H09048). We received 7,041 completed questionnaires, giving a 14.5% response rate.

Variables and measurements
The questionnaire included socio-demographic and gambling behaviour questions. Socio-demographic questions included sex, age, education, workforce status, weekly income-bracket, Indigenous status, household composition, occupation, and residency status. Behavioural questions included name of the respondent's most frequently visited gambling venue (selected from a list), travel mode on last visit, travel-party composition on last visit, standard alcoholic drinks consumed on last visit, and the PGSI. On this scale a score of 8 or above represents a problem gambler, 3-7 a moderate risk gambler, 1-2 a low risk gambler and 0 a non-problem gambler. Occupation was coded using the Australian and New Zealand Standard Classification of Occupations and classified as blue collar or white collar.

In addition, each questionnaire included a unique identifier which enabled us to locate the address to which it was delivered. This enabled us to calculate the distance between each respondent's residential address and their most frequently visited gambling venue. The level of residential socioeconomic disadvantage of the local neighbourhood of the respondent was recorded using the standard Socio-Economic Indexes for Areas Index of Relative Socio-economic Advantage and Disadvantage (SEIFA IRSAD) extracted at census collector district level.

We used recursive partitioning to identify specific gambler subgroups defined through the interaction of the range of predictor variables. Recursive partitioning provides a non-parametric framework for automatically splitting a dataset into progressively smaller groups with increasingly homogenous values of an outcome variable such as the PGSI.

Results
Figure 1 (next page) sets out the results of our analysis in a tree diagram. The problem gambling risk factors highlighted by our analysis (all the variables that are circled in Figure 1) were: visitation party, visitation of a venue with a large number of electronic gambling machines (EGMs) or ‘pokies’, consumption of alcohol on the last visit, being Indigenous, household structure, travel mode, occupational skill level, and age. Variables that are circled towards the top of the tree are more important predictors of gambling risk.

It is these variables in particular combinations that characterize the various subgroups of problem gamblers. One reads the tree in Figure 1 through a series of IF-THEN rules. For example, following the tree along the far left from node 1 (all respondents), if someone visited alone (node 2), visited a venue with less than 45 EGMs (node 3), and were Indigenous, then they ended up in node 4. The boxplot at the end of the node presents the distribution of problem gambling scores for that subgroup (see Figure 2). In the case of node 4 the median PGSI is 4, indicating a moderate- to high-risk group, shown by the heavy grey line in the plot. The box-plot is divided into four groups
containing 25% of responses. The dashed 1st quartile represents the range of the top 25% of PGSI scores (PGSI = 7–13), the grey 2nd quartile representing the range of PGSI scores of the following 25% of respondents (PGSI = 4–6), and so on. No representation of the 4th quartile is visible because all respondents in this quartile had PGSI scores of zero. Outliers (respondents with PGSI = 18, 22 and 25) are shown as dots.

The **riskiest group** detected, with a very high median PGSI of 7 (node 9), included respondents who visited a venue alone (node 2), went to a venue with more than 45 EGMs (i.e. a casino; node 8), and travelled either by taxi, foot or a lift with someone else (node 9). **Other subgroups at elevated risk** of problem gambling were Indigenous people who visited smaller venues alone (node 4), and younger people who travelled alone to larger venues in their own vehicle, by bus or by bicycle (node 11).

The **lowest risk group** detected, with a median PGSI of 0 (node 18), included people who visited with others (node 13), consumed less than 5 alcoholic drinks (node 14), were non-Indigenous (node 16), visited a venue other than the Darwin casino (node 17), and were skilled at bachelor degree or equivalent level (node 18).

In general, recursive partitioning enabled the identification of gambler subgroups with particularly high or low PGSI scores. While the results are limited to the NT and the variables selected, we expect that there is considerable scope for the broader adoption of these methods for gambler subgroup identification. Recursive partitioning of large state and national prevalence studies may well yield new insights into gambling subgroups in those domains, without the need to collect new data.