Estimating the proportion of prisoners in England and Wales who are ex-Armed Forces; a data matching exercise carried out by the MOD in collaboration with the MoJ

EXECUTIVE SUMMARY

1. The aim of this study was to estimate the proportion of prisoners in England and Wales who are ex-Armed Forces by matching administrative datasets held by the MOD and the MoJ. A database of all prisoners in England and Wales on 06/11/09 was matched against a database of Service leavers (Regulars only, going back as far as records are available) using name and date of birth. A match on date of birth and surname (allowing for minor misspelling and truncated surnames, and rejecting matches with differences in forename or middle name) gives 2,207 matches and an estimate of 3%. These figures compare with previous estimates of 6%, 4% and 5% respectively based on surveys carried out by the Home Office in 2001, 2003 and 2004. Although this is the most comprehensive study to date, it is dependent on the quality and completeness of the administrative databases used. In particular, this study will not have captured those who left Service before 1979 (Navy), 1972 (Army) and 1968 (RAF), and so is a lower estimate. Further work is planned during 2010 to estimate the impact of excluding Service leavers prior to the aforementioned dates. We do not expect that a revised estimate would exceed 4%.

BACKGROUND

2. The Home Office conducted surveys of 2,000 nationally representative prisoners at the point of release in 2001, 2003 and 2004. The proportion with Armed Forces experience was reported to be 6%, 4% and 5% respectively.

3. In 2008 the National Association of Probation Officers (NAPO) published findings based on research at HMP Dartmoor and a survey at ten jails by the Veterans in Prison support group in 2007[1]. They concluded that more than 8,000 veterans are currently in prison, many of whom served in Iraq or Afghanistan. Convictions primarily involved violence, within a short period following discharge. The majority of cases were drug- or alcohol-related. These findings suggested that approximately 9% of the England and Wales prisoner population are ex-military.

4. The MOD and MoJ said in response that they planned to carry out further surveys soon to ‘ensure we have up-to-date figures that will help us better target the help we provide for veterans in prison’[2]. The purpose of this collaborative project between the MOD and the MoJ is to determine the proportion of England and Wales prisoners at a certain point in time who are ex-Service personnel.

METHODS

5. This is a cross-sectional study linking a snapshot of all those in prisons in England and Wales with a database compiled by Defence Analytical Services and Advice (DASA) of all discharged military personnel (Regulars only), in order to estimate the proportion of prisoners who are ex-Service.

Databases

6. The MoJ database contains names, date of birth, gender, foreign national status, offence group and date of entry into prison of all remand and sentenced prisoners aged 18 years and over in England and Wales on 06/11/09. The database includes 81,071 records with a unique name and date of birth. Names and dates of birth were standardised (i.e. spaces and hyphens removed) in preparation for matching.

7. The MOD database of Regular Service leavers has been compiled from several sources. It is based on archived data from Service Personnel and Veterans Agency (SPVA) up to 1996 and on ARES (DASA's manpower analysis reporting system) data for 1996 to 2009. The archive data go back to 1972 for the Army, 1968 for the RAF, but only 1993 for the Navy. ARES outflow data,which is available from 1990 onwards, has therefore been used for 1990-1992 inclusive for the Navy. To ensure the database is as complete as possible, further records have been added from DASA's exit data and MEDICS (DASA's database of medical records). The combined dataset contains 1,441,416 records from 1979 (Navy), 1972 (Army) and 1968 (RAF).

8. Data cleaning included the separation of multiple names in one field to facilitate matching, and standardisation of names. Where name or date of birth were missing or obviously erroneous, the records were checked against other MOD sources and completed wherever possible. Otherwise these records were deleted. The total number of records containing both surname and date of birth (and therefore suitable for matching) is 1,327,411. Of these, 68% had a forename and a further 27% had an initial, leaving 6% with missing forename. 42% had a middle name and a further 29% had a middle initial, leaving 29% with missing middle name.

9. Out of these 1,327,411 records, 1,266,648 had unique Service numbers. Ascertainment of Service leavers was assessed by comparing the total number of MOD records with unique Service numbers, by Service, with that expected based on Statement of Defence Estimates (1975 to 1995) and UK Defence Statistics (1996 onwards). This serves as an approximate check only, since we know that a serving individual may be assigned a new Service number for a change of role (e.g. promotion) or a new period of Service and that conversely, the UKDS/Defence Estimate figures will count more than one discharge for individuals who have more than one period of Service. Also, 8% of RAF records had no exit date so could not be included in this comparison. However, for each Service the difference in total number of discharges was 4% or less.

Matching

10. The variables available for matching were surname, forename, middle name, date of birth and sex. Matching strategies of varying stringency were employed by DASA Corporate Systems to estimate the proportion of prisoners who are ex-Service.

11. The matching was divided into several stages. For example, the strictest match required date of birth, surname, forename, middle name and sex to match. Progressively more relaxed matches were then carried out and elements of fuzzy matching were employed to allow for inaccuracies in the name information that might result in true matches being missed. For example, string searches were employed and the first four characters of surnames were matched where truncation of surnames had occurred. Note that sex is not included in all matches because, although it is generally an accurately recorded variable, it has sometimes been found on MOD data sources to be missing or wrongly recorded.

12. This matching strategy resulted in multiple matches i.e. each record could be matched to more than one record in the other database. Possible matches were inspected manually before being accepted or rejected.

RESULTS

13. A strict match on surname, date of birth and initial of forename matched 2.5% of prisoner records to a Service record. A more relaxed match, allowing for misspelling and truncation of surnames (but rejecting matches where there were discrepancies between forename or middle name) suggested that 2,207 prisoners (2.7%) are ex-military.

DISCUSSION

14. This work represents an important addition to current knowledge about ex-Armed Forces personnel in prisons in England and Wales. Previous estimates have been based on relatively small-scale surveys (Home Office), on a single prison (Dartmoor), or have been extrapolated from surveys within a number of selected prisons (Veterans in Prison). All of these methods have their limitations, as does the current study. This matching exercise is the first attempt at a comprehensive study of all prisons in England and Wales, based on administrative data sources. It suggests that the percentage of prisoners in England and Wales who are ex-Armed Forces is approximately 3% (based on 2,207 matches). Since the MOD database used for matching excludes those who left Service before 1979 (Navy), 1972 (Army) and 1968 (RAF), this is likely to be an underestimate. Further work will seek to quantify the likely effect of the incompleteness of the MOD database, although we do not expect that a revised estimate would exceed 4%.

Limitations

15. Data linkage can result in two types of errors: false positive matches and false negative matches. A false positive match is where two records are linked together, when in reality they are not the same person. A false negative match is where two records are not linked together, when they do in fact belong to the same person. Generally there is a trade-off between the two types of errors since, for example, reducing the rate of false positives may increase the rate of false negatives. Assessing the frequency of false positive and false negative matches in this study would require a sample of positive and negative matches to be followed up and questioned as to previous military Service – this presents practical and ethical difficulties and has not been attempted at this stage. We can however outline some potential reasons for erroneous matches.

16. Due to the limited fields available for matching, false positive matches are possible. For example, when the same name appears in both databases alongside the same date of birth this will be counted as a positive match. But, particularly for common names, there is a risk that the records do not actually relate to the same individual.

17. There are several possible reasons that false negative matches may arise. Firstly, because the MOD database of Service leavers is not entirely comprehensive, and does not capture those who discharged before 1979 (Navy), 1972 (Army) and 1968 (RAF), this study may have failed to correctly identify all prisoners who are ex-Service personnel. However, work so far has suggested that convictions are likely to occur soon after discharge – this is likely to reduce the likely impact of the incompleteness of earlier records in the Service leavers database. Secondly, it is expected that matching was less successful for women due to name changes. Although all recorded surnames whilst in UK Armed Forces are included in the MOD dataset, it is possible that after leaving Service a female changed her name then entered prison, in which case she would not be identified in this study. Finally, matches were only accepted if date of birth was an exact match. As a result any individuals whose date of birth was incorrectly recorded on either the MoJ or MOD datasets would contribute to false negative matches. Typical transcription errors in date of birth arise when day and month have been transposed, and when two digits for year are transposed.

18. There are some other general errors in linking databases on date of birth and names that may have affected our results. Problems with names include: variation in spellings, use of nicknames, Anglicisation of foreign names, truncation or abbreviation of names and use of compound names. For some ethnic groups, there can be many surnames and the order of their use varies. Concatenation of the birth surname and the marriage or partnership name into a compound (or hyphenated) name is common, so both parts are required for linking purposes but may not be recorded on both databases.

Recommendations

19. When interpreting the results presented here, it should be borne in mind that this data matching exercise is based on two administrative data sources, the purpose of which is not research. The MOD database in particular represents the best available information at the current time about Service leavers, but we acknowledge it is neither complete nor entirely accurate.

20. In due course it may be possible to quantify the likely effect of the incompleteness of the MOD database, and make an upwardly revised estimate. Note that this is unlikely to increase the estimated proportion of prisoners who are ex-military to more than 4%.

21. The next stage of this project will be to describe the ex-Service prison population in terms of demographic and Service variables, such as Service branch, age, sex, rank and time since discharge. The type of crime committed will also be described. These analyses will enable the MoJ to inform policy and ensure that resources are targeted at appropriate groups of offenders, as well as guiding future research.

References

[1] Ex-Armed Forces Personnel and the Criminal Justice System. NAPO August 2008
[2] Jails 'hold 8,500 ex-servicemen'. BBC Wales News 30/08/08