



Northern  
Ireland  
Office

Statistics and Research Branch

# Predicting Reconviction Rates in Northern Ireland

Research and Statistical Bulletin 7/2005

Brian Francis, Juliet Harman, Leslie Humphreys



# **Predicting Reconviction Rates in Northern Ireland**

**Brian Francis, Juliet Harman, Leslie Humphreys**

**Centre for Applied Statistics,  
Lancaster University**

**May 2005**

**NIO Research and Statistical Bulletin 7/2005**

## Table of contents

Glossary .....	3
1. Introduction .....	4
2. The data .....	4
2.1 Custodial sentence dataset .....	4
2.2 Community sentence dataset .....	4
3. The Offender Group Reconviction Scale (OGRS) .....	4
4. Variables available in Northern Ireland .....	5
5. Statistical Analysis .....	5
5.1 Custodial sentence dataset .....	5
5.2 Community sentence dataset .....	6
6. The preferred models .....	6
6.1 Custodial sentence dataset .....	6
6.2 Community sentence dataset .....	7
7. Predictions .....	7
7.1 Calculation of logit scores .....	7
8. Confidence intervals .....	8
9. Conclusions .....	8
10. Recommendations for further work .....	8
11. References .....	8
Appendix A .....	9
Appendix B .....	9

## Glossary

**Backward elimination technique** - this method begins by entering all the explanatory variables specified into the model. At each step, the least significant explanatory variable is removed from the model until all of the remaining explanatory variables have a statistically significant contribution to the model.

**Binary dependent variable** - a variable which can take only two values, (Yes/No; success/fail)

**Bootstrap (re-sampling)** - a technique commonly used to estimate confidence intervals by repeatedly resampling from the observed data.

**Categorical variable** - a variable is said to be categorical if the values/observations belonging to it can be sorted according to category. Each value is chosen from a set of non-overlapping categories. For example people have the characteristic of 'gender' with categories 'male' and 'female'.

**Confidence levels** - measure the likelihood that the true value of a parameter is within a certain distance of the results of a sample analysis. A 95% confidence level is commonly used. This means that 95% of the time, the true value for the population will lie within the the confidence interval.

**Confidence interval** - a confidence interval gives an estimated range of values which is likely to include an unknown population parameter, the estimated range being calculated from a given set of sample data. Confidence intervals are more informative than the simple results of hypothesis tests since they provide a range of plausible values for the unknown parameter.

**Continuous variable** - a variable is said to be continuous if the values/observations belonging to it may take on any value within a finite or infinite interval. You can count, order and measure continuous data. For example height, weight, temperature, the amount of sugar in an orange, the time required to run a mile.

**Indicator variable** - is a variable that takes the value one if a condition is true, and zero if it is false.

**Logistic Regression** - is a statistical technique used to model binary dependent variables (that is those which take only two values (yes/no; success/failure)).

**Logit transformation** -the logit (or logistic) transformation is a transformation of a binary dependent variable.

**Functional form** – is the mathematical relationship (e.g. square root) among the variables in a model.

**Parameter**- is a value, usually unknown (and which therefore has to be estimated), used to represent a certain population characteristic.

**Reconviction rate** – the percentage of offenders in a given year who are reconvicted, for any offence, within a specified period (two years in the case of Northern Ireland studies) from the date of their discharge from custody or receipt of their non-custodial disposal.

**Response variable** – a variable whose values are to be explained by other variables.

**Standard list offences** – these include all 'indictable only' and 'triable either way' offences and some of the more serious 'summary' offences such as common assault and driving whilst disqualified or under the influence of drugs or alcohol.

**Transformation** - transformations allow us to change all the values of a variable by using some mathematical operation, for example, we can change a number, group of numbers, or an equation by multiplying or dividing by a constant or taking the square root. A transformation to linearity is a transformation of a response variable, or independent variable, or both, which produces an approximate linear relationship between the variables.

**Wald test p-value** - this statistic is a test of significance of the regression coefficient.

# 1. Introduction

The aim of this Bulletin is to inform policy colleagues in their work towards the achievement of Target 3 of the Northern Ireland Office Public Service Agreement which looks at reducing the rate of reconviction in Northern Ireland.

Recent data was used to develop statistical models for predicting the rates of reconviction within two years of receiving a community sentence or release from custody in Northern Ireland. Observed rates can then be compared with the predicted rates.

The models are based on the Offender Group Reconviction Scale (OGRS) developed by Copas and Marshall (1998), and its later modification by Taylor (1999). They assume that the probability of reconviction can be associated with a small number of influential factors relating to the demographic characteristics and criminal history of the offender.

Using data for male and female offenders sentenced to a community penalty (Probation Order, Community Service Order, Combination Order) or released from custody in Northern Ireland in 1998, 1999 and 2000 whose criminal histories had been traced for two years, the significant variables associated with reconviction within this period for both types of disposal can be identified.

The models were then used to predict reconviction rates and confidence intervals for offenders who had received community sentences or who had been released from custody. Predicted rates can then be compared with the observed rates.

The predictions assume that past and current offender populations behave in similar ways and **do not include** an estimate of social change in Northern Ireland or any changes in sentencing practices and offender treatments.

## 2. The data

The system for generating reconviction rates has been developed by statisticians out-posted to the Northern Ireland Office from the Northern Ireland Statistics and Research Agency. This system is based on data supplied by the Police Service of Northern Ireland (PSNI) from their Integrated Crime Information System (ICIS). PSNI are responsible for the structure and accuracy of the data held on ICIS.

### 2.1 Custodial sentence dataset

The data used for calibrating the statistical model was the total population of offenders who had been released from custody in 1998, 1999 and 2000 (2,844 cases<sup>1</sup>). Of the 2,844 cases for which reconviction data was available, 49.2% were reconvicted within two years. The

<sup>1</sup>If an offender is convicted in more than one of these years s/he will be included more than once. So an offender convicted once in 1998 and once in 1999 will be included twice.

reconviction rate for offenders released from Immediate Custody (IC) was 49.7%, the reconviction rate for those released from Custody Probation Order (CPO) was 41.0%, and the reconviction rate for those released from the Juvenile Justice Centre (JJC) was 83.5%.

### 2.2 Community sentence dataset

The data used for calibrating the model contained 4,792<sup>2</sup> offenders who had received a community penalty (Probation Order, Community Service Order, Combination Order, Attendance Centre Order) in 1998, 1999 and 2000. Of the 4,792 cases for which reconviction data was available, 44.7% were reconvicted within two years. The reconviction rate for offenders sentenced to a Probation Order was 45.5%, the reconviction rate for those sentenced to a Community Service Order was 42.2%, the reconviction rate for those sentenced to a Combination Order was 49.1%, and the reconviction rate for those sentenced to an Attendance Centre Order was 58.8%.

Pseudo-reconvictions are not included in the datasets. Pseudo-reconvictions are convictions that were 'in the pipeline' i.e. for offences committed prior to the date of the baseline conviction. Their exclusion ensures that they do not contribute to generating artificially high reconviction rates and artificially short reconviction intervals.

## 3. The Offender Group Reconviction Scale (OGRS)

The Offender Group Reconviction Scale (Copas and Marshall, 1998) is based on logistic regression analysis using large samples of offenders. The initial exploratory analyses included approximately 14,000 offenders who had received a community sentence or had been released from prison in 1987 in England and Wales, whose criminal histories were traced on the Offenders Index for two years. The model was refined using 1990 data, using over 4,000 cases. Only standard list offences as defined by the Home Office in 1990 were included.

A variable indicating whether or not the offender had been reconvicted was the response variable in the model. The variables shown in Table 1 were found to be good predictors of reconviction within two years:

Table 1: Predictors used in the OGRS

◆ Offender's age
◆ Offender's gender
◆ Number of youth custody sentences
◆ Total number of court appearances leading to a conviction
◆ Time in years since first conviction
◆ Offence group for baseline offence (9 categories)

Copas and Marshall combined the total number of court appearances leading to a conviction (*n*) and the time since the first conviction (*t*) in a conviction rate variable, which,

<sup>2</sup>As was the case with the custodial dataset, it is possible for offenders to appear in the non-custodial data more than once if s/he was convicted in more than one year between 1998 and 2000.

after exploratory analyses, was defined as the square-root of  $(n/(t+5))$ . The transformation seemed appropriate in order to include individuals for whom the current conviction was their first. This variable is usually referred to as the Copas rate.

A revised version of OGRS is now used (Taylor, 1999), in which the current offence group is divided into 27 categories, and indicator variables for a history of burglary and a history of breach are also included. This model is based on a sample of over 20,000 offenders from the 1995 cohort.

## 4. Variables available in Northern Ireland

The following variables were extracted from the ICIS system and used as potential predictors in the statistical analysis. A slightly different set of variables was available for the custodial offenders and for the community sentence offenders.

Table 2a: Prediction variables available for calibration

Variable definition	Custodial data	Community sentence data
Type of disposal for baseline offence	y	y
Age at discharge for baseline offence	y	-
Age at conviction for baseline offence	-	y
Sex of offender	y	y
Age at first conviction	y	y
Age at first custodial sentence	y	y
Indicator of any convictions previous to baseline	y	y
Number of convictions before baseline offence	y	y
Number of convictions previous to baseline in NINE offence categories (violence against person, sex, burglary, robbery, etc.)	y	y
Length of custodial sentence in days	y	-
Baseline offence scheduled or not	y	y
Baseline offence class	y	y
Baseline offence qualifier (indictable, summary, motoring)	-	y
Indicator of custody previous to baseline	y	y

Table 2b: Response variables available for calibration

Variable definition	Custodial data	Community sentence data
Reconviction offence qualifier (indictable, summary, motoring)	y	y
Indicator of Reconviction within two years of baseline	y	y
Reconviction interval in days	y	y
Reconviction scheduled or not	y	y
Reconviction offence class	y	y
Age at first custodial sentence	y	y

## 5. Statistical Analysis

We proceeded by building separate statistical models for the probability of reconviction for those finishing a custodial sentence, and for those sentenced for a community disposal. Logistic regression analysis was used to build each of these models, with reconviction within two years as the response variable and a subset of variables from Table 2a which represented standard OGRS variables as predictors. For the custodial dataset, length of prison sentence for the baseline offence, type of custody and a variable indicating whether or not the baseline offence was scheduled were also used.

The significant predictor variables were found by using a backward elimination technique, further details of the method are given in appendix A.

### 5.1 Custodial sentence dataset

The following explanatory variables were used in the backward elimination, with some recoding of the original data as suggested by preliminary analyses:

- ◆ age at first conviction
- ◆ age at discharge
- ◆ sex of offender
- ◆ baseline offence class (11 categories)
- ◆ conviction rate (cf. Copas rate) can be calculated from total number of convictions and length of criminal history.
- ◆ sentence length
- ◆ whether offender had a previous principal conviction for each of nine separate offence types.
- ◆ offender disposal type (IC, CPO, JJC)
- ◆ whether the offender had a custodial sentence while a juvenile (under 17 years old)
- ◆ whether the baseline offence was scheduled
- ◆ total number of prior conviction occasions.
- ◆ whether the offender had a previous custodial sentence.

### The conviction rate variable

Exploratory analysis showed the conviction rate variable to be the most important predictor. To obtain the optimal functional form for this variable, the likelihood profile of a range of functions  $f$  and constants  $k$  was investigated for  $rate=f(n/(t+k))$ . In this equation  $n$  is the total number of convictions, including the baseline conviction;  $t$  is the number of years between first conviction and release from prison, and  $k$  is a constant to be determined.

The square root function used by Copas gave the highest log-likelihood with seven variables in addition to rate included in the model. As the log-likelihood was found not to be very sensitive to the choice of  $k$ ,  $k=5$  as used in the Copas rate was adopted for this dataset.

### Recoding variables

Length of sentence, age at discharge and age at first conviction were treated both as continuous and categorical variables. Preliminary analysis suggested that models which categorised these variables fitted better than those using continuous variables. Various categorical schemes were chosen, and the scheme which offered the best fit chosen.

### 5.2 Community sentence dataset

The following explanatory variables were used in the backward elimination, with some recoding, as suggested by preliminary analyses:

- ◆ age at first conviction
- ◆ age at baseline conviction
- ◆ sex of offender
- ◆ baseline offence class (11 categories)
- ◆ conviction rate (cf. Copas rate) can be calculated from total number of convictions and length of criminal history.
- ◆ whether offender had a previous principal conviction for each of nine separate offence types.
- ◆ Any previous convictions
- ◆ offender disposal type
- ◆ whether the offender had a custodial sentence while a juvenile (under 17 years old)
- ◆ whether the baseline offence was scheduled
- ◆ total number of prior conviction occasions.
- ◆ whether the offender had a previous custodial sentence.

### The conviction rate variable

An investigation of the likelihood profile (see above) of a range of functions  $f$  and constants  $k$  in  $rate=f(n/(t+k))$  suggested  $rate=(n/(t+5))$  to be most appropriate for this dataset, with the identity and square root functions giving the highest log-likelihood. Again, the square root transformation was chosen.

### Recoding variables

Age at baseline conviction and age at first conviction were treated both as continuous and categorical variables. As in section 5.1, preliminary analysis suggested that models which categorised these variables fitted better than those using continuous variables. Various categorical schemes were chosen, and the scheme which offered the best fit

chosen. In particular, age at baseline conviction had a finer classification than the variable age of discharge used in the custodial analysis.

## 6. The preferred models

The constant term in the following models refers to an offender in the lowest category for each explanatory variable. For offenders in other categories, the appropriate parameter estimates have to be added as indicated below.

### 6.1 Custodial sentence dataset

Table 3: Parameter estimates for preferred model (Custodial sentence data)

Explanatory Variable	Parameter estimate	Standard error of estimate	Wald test p-value
Constant	-0.349	0.337	0.300
COPAS RATE <sup>3</sup>	2.124	0.196	<0.001
BOFFENCE(VAP)	0.0	-	-
BOFFENCE(SEX)	-1.060	0.312	0.001
BOFFENCE(BURGLARY)	0.218	0.177	0.217
BOFFENCE(ROBBERY)	0.015	0.174	0.930
BOFFENCE(THEFT)	-0.021	0.157	0.895
BOFFENCE(FRAUD/FORG)	-0.333	0.265	0.209
BOFFENCE(CRIM DAM)	0.444	0.192	0.021
BOFFENCE(OAS)	0.427	0.270	0.113
BOFFENCE(OTHER)	0.211	0.194	0.277
BOFFENCE(MOTORING)	-0.2734	0.173	0.113
BOFFENCE(DRUGS)	-0.079	0.168	0.637
AGEDISCH(10-17)	0.0	-	-
AGEDISCH(18-20)	-0.809	0.308	0.009
AGEDISCH(21-24)	-1.363	0.307	<0.001
AGEDISCH(25-29)	-1.459	0.310	<0.001
AGEDISCH(30 and over)	-1.615	0.305	<0.001
AGE1CONV(10-13)	0.0	-	-
AGE1CONV(14-17)	-0.081	0.143	0.571
AGE1CONV (18-24)	-0.203	0.153	0.186
AGE1CONV (25 and over)	-0.768	0.256	0.003
ANYVIOLENCE	-0.254	0.133	0.012
ANYBURGLARY	0.418	0.106	<0.001
SENTENCE (under two years)	0.0	-	-
SENTENCE (2 years to under 4 years )	-0.160	0.133	0.229
SENTENCE (4-30 years)	-0.377	0.206	0.068
SENTENCE (life)	-2.013	0.973	0.049

<sup>3</sup> rate =  $\sqrt{n / (t+5)}$  (See section 5.1)

Table 4: Variable definitions for preferred model (Custodial sentence data)

Variable name	Type	Definition
COPAS RATE	Continuous	Copas rate of conviction between first conviction and release from custody <sup>3</sup>
BOFFENCE	Categorical	Principal baseline offence leading to the custodial sentence
AGEDISCH	Categorical	Age at discharge from baseline offence
ANYVIOLENCE	Indicator	Violence against the person conviction prior to baseline
ANYBURGLARY	Indicator	Burglary conviction previous to baseline
SENTENCE	Categorical	Baseline sentence length
AGE1CONV	Categorical	Age at first conviction

<sup>3</sup> rate =  $\sqrt{(n / (t+5))}$  (See section 5.1)

## 6.2 Community sentence dataset

Table 5: Parameter estimates for preferred model (Community sentence data)

Explanatory Variable	Parameter estimate	Standard error of estimate	Wald test p-value
Constant	0.315	0.228	-
COPAS RATE <sup>4</sup>	1.901	0.110	<0.001
AGE(10-14)	0.0	-	-
AGE(15-17)	-0.331	0.177	0.06
AGE(18-19)	-0.747	0.184	<0.001
AGE(20-21)	-1.319	0.193	<0.001
AGE(22-24)	-1.445	0.193	<0.001
AGE(25-29)	-1.867	0.192	<0.001
AGE(30-39)	-2.038	0.187	<0.001
AGE(40 or over)	-2.110	0.199	<0.001
FEMALE OFFENDER	-0.491	0.098	<0.001
BOFFENCE(VAP)	0.0	-	-
BOFFENCE(SEX)	-0.486	0.311	0.12
BOFFENCE(BURGLARY)	0.183	0.151	0.23
BOFFENCE(ROBBERY)	0.236	0.384	0.54
BOFFENCE(THEFT)	0.096	0.123	0.45
BOFFENCE(FRAUD/FORG)	0.185	0.187	0.32
BOFFENCE(CRIM DAM)	-0.058	0.139	0.68
BOFFENCE(OAS)	0.069	0.264	0.79
BOFFENCE(DRUGS)	0.182	0.242	0.78
BOFFENCE(OTHER INDICT)	0.074	0.261	0.45
BOFFENCE(SUMMARY)	-0.204	0.138	0.14
BOFFENCE(MOTORING)	-0.227	0.150	0.13
BSCHEDULED	-0.432	0.144	0.003

Table 6: Variable definitions for preferred model (Community sentence data)

Variable name	Type	Definition
COPAS RATE	Continuous	Copas rate of conviction between first conviction and baseline conviction <sup>4</sup>
AGE	Categorical	Age at baseline conviction
BOFFENCE	Categorical	Baseline principal offence category
BSCHEDULED	Indicator	Scheduled principal offence (1=yes, 0=no)
FEMALE OFFENDER	Indicator	Whether offender is female (1=yes, 0=no)

<sup>4</sup> rate =  $\sqrt{(n / (t+5))}$  (See section 5.2)

## 7. Predictions

Reconviction rates for new populations of offenders can be predicted provided the variables in the above models are available for each offender. The predictions assume that past and current offender populations behave in similar ways and therefore do not include any changes in sentencing practices and offender treatments.

The parameter estimates in the above models were used to calculate the fitted probabilities of reconviction. First a logit score was calculated for each offender as described below. This was then converted to a probability of reconviction for each offender. The probabilities for the whole population (custodial or non-custodial) were added and converted to a percentage, to give the predicted reconviction rate.

### 7.1 Calculation of logit scores

The parameter estimates in the preferred model were used to calculate the predicted logit score for each offender as described here for *the custodial sentence dataset (see Table 3)*.

A similar procedure was followed for the community sentence dataset.

- a) start from the constant of -0.349
- b) add COPAS RATE multiplied by 2.124
- c) if the baseline offence was a violence offence add 0; if a sexual offence subtract 1.060; if it was burglary add 0.218; if robbery, add 0.015; if theft, subtract 0.021; if fraud or forgery subtract -0.333; if criminal damage, add 0.444; if offences against the state, add 0.427; if motoring subtract 0.274; if drugs, subtract 0.079; and if an 'other' offence add 0.211.

- d) if the age at discharge from custody for the baseline offence was between 10 and 17, then add 0;  
if between 18 and 20, then subtract 0.809;  
if between 21 and 24, then subtract -1.36;  
if between 25 and 29 then subtract -1.459;  
and if aged 30 or over, subtract -1.62;
- e) if the offender's age at first conviction was 13 or less, then add 0;  
if between 14 and 17, then subtract 0.081;  
if between 18 and 24 then subtract 0.203;  
and if 25 or over, then subtract 0.768.
- f) if the offender had a previous burglary conviction add 0.418, otherwise add 0
- g) if the offender had a previous violence conviction subtract 0.254, otherwise add 0.
- h) if the sentence length for the baseline offence was under two years then add 0;  
if it was between 2 years and under 4 years, then subtract 0.160;  
if between 4 years and 30 years then subtract 0.377;  
and if a life sentence then subtract 2.013

The fitted probability of  
reconviction within two years =  $\frac{\exp(\text{logit score})}{(1 + \exp(\text{logit score}))}$

The expected reconviction rates were found by summing the fitted probabilities for all offenders in each category, and were converted to percentages. It should be noted that predicted values are not presented in this Bulletin as it deals specifically with the methodology of developing the models.

## 8. Confidence intervals

The 95% and 99% confidence intervals for the predicted rates were found by bootstrap (re-sampling) methods (Efron and Tibshirani, 1993). See Appendix B for further details of this method.

Confidence intervals for the predicted rates were obtained for each sentence type, and also for different types of discharge or disposal.

## 9. Conclusions

This analysis has produced predictions for two-year reconviction rates for offenders released from custody or sentenced to a community sentence in 2001-2003, using data observed in 1998, 1999 and 2000. We used three years of data to make predictions as this provided a great improvement in the power of the model, and enabled us to include a wide variety of predictor variables and to develop models which were similar in nature to the OGRS models of the Home Office.

Two points need to be made. Firstly, the predictions assume that the nature of offending and the behaviour of offenders in 2001-2003 remain similar to offenders observed in 1998-2000. Thus, the predictions will not take into account social and demographic change, legislation, crime reduction measures and other initiatives brought in to combat crime. Social and demographic change might well affect reconviction rates, whereas government intervention would hope to reduce crime.

The second point is that the observed reconviction rates declined over the three year period from 1998-2000 for those receiving community sentences. The statistical model does not predict this decline from the criminological variables used in the model, and it is likely that this decline is caused by other factors.

## 10. Recommendations for further work

It is essential that these current predictor models be reviewed periodically in the future to take account, where possible, of forthcoming changes in legislation and as new disposals available to the courts are brought forward.

When sufficient time has passed to allow the above changes feed through the criminal justice system the reconviction predictor model should be remodelled based on the new data.

## 11. References

- Copas, J. and Marshall, P. (1998), The offender group reconviction scale: a statistical reconviction score for use by probation officers, *Applied Statistics*, 47, 1, pp.159-171
- Efron, B. and Tibshirani, R. J. (1993), *An introduction to bootstrap*, Chapman and Hall
- Taylor, R. (1999), *Predicting reconvictions for sexual and violent offences using the revised offender group reconviction scale*. Home Office Research Findings No.104. Home Office, London

## Appendix A

Logistic regression is a statistical technique to model binary dependent variables (that is those which take only two values (yes/no; success/failure)). In this report, the variable is whether or not an offender has been convicted within two years of sentence (for community sentence offenders) or release from custody (for those receiving custodial sentences).

The method builds a model for the probability ( $P_i$ ) of reconviction within two years for any offender  $i$  in terms of a set of explanatory or predictor variables. Formally, we have the relationship

$$\text{logit}(P_i) = \log \left( \frac{P_i}{1 - P_i} \right) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \sum_{a=1}^A \beta_i^a X_{ai} + \sum_b^B \beta_i^b X_{bi} + \dots$$

where  $X_1, X_2 \dots$  are continuous variables, and  $X_{1i}, X_{2i}$  are the values of that variable for offender  $i$ ; and where  $X_A, X_B \dots$  are categorical variables with A, B,... levels and where  $X_{ai}, a=2..A, X_{bi}, b=2..B, \dots$  are dummy variables formed from these categorical variables, with the first category as the baseline category. The betas are unknown parameters to be estimated.

The model was fitted using maximum likelihood, and estimates of the betas and their standard errors are obtained. These betas and their standard errors are presented in Tables 3 and 5 of the main report.

Model selection was carried out by backward elimination. A model with all predictor variables is first fitted, and the value of minus twice the log-likelihood (or  $-2 \log l$ ) of the model is calculated. Each variable is then removed from the model, and  $-2 \log l$  values calculated. For any omitted variable, differences in  $-2 \log l$  are chi-squared distributed with  $k$  degrees of freedom if the beta parameter or parameters for that variable can be set to zero. For continuous variables,  $k=1$  and for categorical variables,  $k$  is equal to the number of categories minus 1. We omit the variable with the largest p-value which is also insignificant at the 5% level, and continue. The process stops when no further variables can be eliminated.

We can use the model to produce estimated logit scores -  $\text{logit}(P_i)$  – and estimated predicted probabilities  $P_i$  for any new case or set of new cases – simply by substituting the values of the predictor variables into the above equation for each new case. Summing the estimated  $P_i$  over the set of new cases provides an estimate of the number of predicted reconvictions in the set.

## Appendix B

The 95% and 99% confidence intervals for the predicted rates were found by bootstrap (re-sampling) methods (Efron and Tibshirani, 1993).

This involved re-sampling the 1998, 1999 and 2000 data used in model calibration 10,000 times, with replacement. The preferred logistic regression models were fitted to each of these new samples, giving a range of parameter estimates and fitted probabilities for each offender. From these fitted probabilities, a predicted reconviction rate for each new sample was calculated. The 95% and 99% confidence intervals were obtained from the distributions of the predicted reconviction rates.

## NOTES

---

## NOTES

---

## Other publications available from NIO Statistics and Research Branch:

### Research & Statistical Series

Report No.11	Evaluation of Youth Justice Agency Community Services	July 2004
Report No.10	Victims' and Witnesses' Views on their Treatment in the Criminal Justice System	February 2004
Report No.9	The detention and questioning of young persons by the police in Northern Ireland K Quinn & J Jackson	November 2003

### Research & Statistical Bulletins

6/2005	Juvenile Reconviction in Northern Ireland 2001 M Decodts	March 2005
5/2005	Experience of Domestic Violence in Northern Ireland: Findings from the 2003/04 Northern Ireland Crime Survey R Freel & E Robinson	March 2005
4/2005	Crime Victimization in Northern Ireland: Findings from the 2003/04 Northern Ireland Crime Survey B French & P Campbell	March 2005
3/2005	Adult Reconviction in Northern Ireland 2001 S McMullan & D Ruddy	February 2005
2/2005	Northern Ireland Statistics on the Operation of the Terrorism Act 2000: January - September 2004 D Lyness & M Carmichael	February 2005
1/2005	Interim Evaluation of the Northern Ireland Youth Conferencing Scheme H Beckett, C Campbell, D O'Mahony, J Jackson & J Doak	January 2005
6/2004	Perceptions of and concern about crime in Northern Ireland: Findings from the 2003/04 Northern Ireland Crime Survey P Campbell & B French	December 2004
5/2004	Public attitudes towards crime and recovery of assets by the Assets Recovery Agency in Northern Ireland: Findings from the January 2004 Northern Omnibus Survey D Ruddy & D Lyness	October 2004
4/2004	Views on Organised Crime in Northern Ireland: Findings from the April 2004 Northern Ireland Omnibus Survey M Wilson & B French	October 2004
3/2004	Northern Ireland Statistics on the Operation of the Terrorism Act 2000: Annual Statistics 2003 D Lyness & M Carmichael	August 2004
2/2004	The Northern Ireland Prison Population in 2003 S McMullan, K Amelin & M Willis	June 2004
1/2004	Experience of Domestic Violence in Northern Ireland: Findings from the 2001 Northern Ireland Crime Survey R Freel & E Robinson	March 2004

### Compendia Publications

A Commentary on Northern Ireland Crime Statistics 2003	June 2004
Digest of Information on the Northern Ireland Criminal Justice System 4	March 2004
The Use of Proxy Measures of Religion (and other Section 75 categories) in relation to the Northern Ireland Criminal Justice System	February 2004

Produced by the Statistics and Research Branch of the Northern Ireland Office.  
For additional copies contact:

Statistics and Research Branch  
Criminal Justice Policy Division  
Massey House  
Stoney Road  
Belfast BT4 3SX

**Telephone: 028 9052 7534**

**Fax: 028 9052 7532**

**Email: [statistics.nio@nics.gov.uk](mailto:statistics.nio@nics.gov.uk)**

This publication and others on related criminal justice areas are also available on the  
Internet at:

**[www.nio.gov.uk](http://www.nio.gov.uk)**

**The views in this report are those of the authors and are not to be attributed to the  
Northern Ireland Office**

Application for reproduction should be made to Statistics and Research Branch,  
Massey House, Stoney Road, Belfast, BT4 3SX.

**Northern Ireland Office**

Criminal Justice Directorate  
Statistics and Research Branch  
Massey House  
Stoney Road  
Belfast  
BT4 3SX

Email: [statistics.nio@nics.gov.uk](mailto:statistics.nio@nics.gov.uk)

Telephone: 028 9052 7534

Facsimile: 028 9052 7532

[www.nio.gov.uk](http://www.nio.gov.uk)

