Developing a reoffending measure using the Police National Computer database

Centre for Applied Statistics
Lancaster University

FINAL REPORT

Principal investigators:

Brian Francis	Centre for Applied Statistics
Keith Soothill	Department of Applied Social Science
Les Humphreys	Centre for Applied Statistics

Contact:

Professor Brian Francis, Centre for Applied Statistics, Fylde College, Lancaster University, LA1 4YF

Phone: (01524) 593061 Fax: (01524) 593429
E-mail: B.Francis@Lancaster.ac.uk

Table of contents

Executive summary

This work aims to develop a reoffending score for assessing the likelihood of an offender reoffending within a certain fixed period of time. Currently, OGRS - the Offenders Group Reconviction Score - is a heavily used measure for assessing reconviction scores of a group or set of offenders. However, dates of conviction can be far removed from the reality of offending. The aim in this report is to develop a new approach which focuses on 'offending' rather than 'conviction' and, hence, to provide a 'reoffending measure' rather than a 'reconviction measure'.

While OGRS has been remarkably successful, there is a need to develop a new measure. There are various reasons - to make offending behaviour rather than convictions as the primary focus of interest; the existing score is somewhat outdated; to confront the problems with the existing OGRS score; to improve the documentation relating to the construction of a measure; and to make the measure more 'user-friendly' for the practitioners.

The construction of the OGRS was based on the Offenders Index (OI) as the data source. This project is concerned with developing a new reoffending score, based on information from the Police National Computer (PNC). The PNC provides the potential for including a wider range of variables that make feasible the transition from a 'convection measure' to a 'reoffending measure'.

Issues relating to the use of the PNC data

The Police National Computer data offers a better coverage of contemporary criminal histories than the Offenders Index. For example, among the advantages are: cautions and warnings are recorded, the database covers Scottish offending as well as England and Wales, and criminal histories are drawn from operational records and will be more reliable than those from the OI.

The report confronts the issues which could potentially affect the development of a new measure. These are: (a) the weeding policy of criminal convictions, (b) the more detailed information on offences, c) the relative lack of historical information through back-record conversion and d) the inclusion of Scottish offences and offences from other police forces.

Developing a new measure

It is argued that a new measure should not be regarded as simply a minor development of OGRS, for the current work provides a genuine re-conceptualisation of the approach to measuring - that is, moving from conviction data to a greater focus on offending behaviour. A fundamental re-naming removes the danger of spurious attempts to compare the outcomes of the two measures. Our proposed name for the new approach - ASPRO (A Score for Proven Reoffending of Offenders) - reflects the increased focus on 'reoffending', rather than on 'reconviction', as the pivotal concept.

In developing a new measure, the report confronts various tasks. First, we define some terminology. Secondly, we specify precisely what is meant by reoffending and, thirdly, we operationalise what predictor variables can be used.

- ***Terminology.*** The main changes relate to the use of reoffending dates (which were unavaiblae on the Offenders Index) rather than reconviction dates as the pivot of the one- and two-year follow-up analysis. However, the PNC is an operational database and we need to focus on *proven* offences. We therefore recommend that "date of *proven* reoffending" is used. We consider all offending which occurs within the two-year or one-year follow-up period, but insist that this offending is "proved" either by a caution (or other police action) or by a court conviction. This means that there needs to be an additional period (which we term a "confirmation period") in which offences which occur towards the end of the two-year (or one year) follow-up period have a chance of being proved and appearing on the PNC database. Furthermore, using PNC data, the term 'conviction' is unnecessarily restrictive, for there is scope also to consider cautions, warnings and reprimands. By broadening the scope of the disposals, we get much nearer to the notion of actual offending behaviour. We use the term "sanction" to refer to cautions, warnings and reprimands as well as convictions.

- ***What is meant by reoffending?*** It is important to define both the follow-up period and the confirmation period. The longer the confirmation period is, then the greater the proportion of offenders who will be proved in court. For this study we use the following as a definition of reconviction/reoffending.

  An offender who has committed a recordable offence within the follow-up period and who has had the offence "proved" within the follow-up period and a confirmation period of 3 months, either by the offender accepting a caution, warning or reprimand, or by being found guilty in a court of law.

- ***The predictor variables.*** In the report a large set of potential candidate variables – either variables which have been used as predictors in OGRS studies, either by the Home Office or by the Northern Ireland Office, or potential new variables suggested by this research - is considered.

The datasets

Two datasets consisting of the following samples were supplied from the PNC:

a) A set of 81744 offenders with non-custodial sanctions in England and Wales in January 2002.
b) A set of 7921 offenders released from custody in January, February or March 2002.

After 'cleaning up' the data for further analysis, the final dataset gave a total of 71519 non-custodial offenders and 7675 custodial offenders.

Constructing the reoffending measure

In constructing the new reoffending measure, the dataset was first divided into two. Around 60% of the sample (the calibration dataset) was used to construct the reoffending score, and the remaining 40% of the data (the validation dataset) was not used in the model fitting but reserved to determine the accuracy of the score.

Models were then fitted to the calibration dataset. Once a final model had been determined, the model was assessed by comparing the fitted probabilities from the model with the actual reoffending measures for the validation sample

For the combined calibration dataset, we fitted a series of logistic regression models, modelling the probability of two-year reoffending as a function of explanatory variables. We repeated the process for one-year reoffending. We adopted a procedure which started with a similar set of variables to those used in the earlier OGRS measures, and adapted the model to improve the fit of the model. We determined that criminal history measures were best summarised through counting sanctions rather than offences, and that categorisation of variables such as age provided a better fit than the continuous equivalents. We also determined that the "Copas rate" should be modified by replacing the square root transformation by a log transformation. The final model included variables for age at sanction, target offence at sanction, criminal history at target sanction, gender, previous juvenile custody, and custody over four years, as well as the revised Copas rate..

A competing model used ordinal regression. This model has the advantage that it provides a method of combining the two logistic regression models, allowing a single score to be used for predicting both one-year and two-year reoffending. Selection of variables produced a model with exactly the same variables as for the earlier logistic regressions.

Assessing the accuracy of the measures.

The models were assessed and compared by using the remaining 40% of offenders not used for the model construction. Two criteria were used - the AUC or area under the ROC curve for the final fitted models, and the fitted percentage of offenders reoffending after one year and after two years (both with a three month confirmation period) with the estimated percentage of offenders predicted by the model. The first criterion can be thought of as a measure of concordance between the predicted fit of the model and the actual reconviction outcome.

The AUC criterion showed that both the logistic regression and the ordinal regression showed excellent concordance, with values around 0.8, and there was little difference between the two models. Examination of the fitted percentages showed that the model was successful in estimating observed reoffending by offence group, and also worked well for male offenders. However, the model exhibited some limitations for predicting female reoffending, and worked better for two-year offending.

Conclusions

The aim of this report has been to develop a new measure to assess the performance of offenders following a sanction. This has involved a reappraisal of the current OGRS measure which has been widely and successfully used for assessing reconviction probabilities for groups of offenders. The new measure focuses on reoffending behaviour rather than the narrow focus on reconviction.

We maintain that this reconceptualisation is so fundamental that the new measure needs a new name, rather than simply being a variant of the OGRS measure. The shortcoming of the OGRS name for this new work is that it misleadingly places reconviction too prominently in the title. We suggest that the name ASPRO (A Score for Proven Reoffending of Offenders) more clearly represents the nature of the work undertaken.

We have presented two possible score sets. The set based on logistic regression provides two separate scores – one for one-year reoffending and one for two-year reoffending. The second set, in contrast, provides a single score which can be used to predict both one and two-year reoffending.

1. Introduction

This work aims to develop a reoffending score for assessing the likelihood of an offender reoffending within a certain fixed period of time. Currently, OGRS - the Offenders Group Reconviction Score - is a heavily used measure for assessing reconviction scores of a group or set of offenders. However, dates of conviction can be far removed from the reality of offending. The aim in this report is to develop a new approach which focuses on 'offending' rather than 'conviction' and, hence, to provide a 'reoffending measure' rather than a 'reconviction measure'. The task is both to explain this reconceptualisation and to describe the processes which lead to a reoffending measure using the Police National Computer database.

1.1 Background to reoffending and reconviction scores

The development of OGRS has an important history. Two versions of the score have been produced – the original was developed by John Copas and Peter Marshall in 1993 (Copas and Marshall, 1998) and a revised version – OGRS2 – was developed in 1998 (Taylor, 1999).

Although originally developed as a *group* measure of reconviction, the OGRS score is now used routinely for assessing an *individual'*s risk of reconviction. Probation staff routinely use OGRS scores together with information from OASys to assess the level of risk of an offender being reconvicted. OGRS scores are also available as part of a MAPPP assessment for dangerous offenders. However, the scores are also used in evaluation studies, and can be used to ensure that control and experimental groups are comparable when comparing interventions in a quasi-experimental context. So, for example, an assessment of the Pathfinder programme in the probation service used OGRS2 as a control measure when comparing experimental and control groups (Rex et al, 2004).

Other uses for OGRS include a prediction of the rate of offending in future years based on the criminal histories of those before the courts - and this provided a baseline rate of reconviction so that targeted improvements in such rates can be assessed.

The work for the existing score used the Offenders Index (OI), which is a court –based recording system. It records the date of court-based convictions for all courts in England and Wales. Convictions are not all recorded – only those involving a standard list offence. Historical information is good – convictions have been recorded since 1963.

1.2 What does OGRS do?

Formally, the OGRS score estimates the two-year reconviction probability for an offender at a certain point in time which is called the target date – either immediately after sentence, or at release from custody. That conviction which resulted in the sentence or custodial period is called the target conviction – the group of offences making up that conviction are called the target offences.

The reconviction is defined as a subsequent conviction in a court for a standard list offence – that is, any indictable offence and the more serious summary offences. The reconviction must take place within two years of the target date. In practice, an additional period is used of about three months – reconvictions are not counted if they occur in this period, but it allows courts sufficient time to place information on the OI dataset. A list of standard list offences can be found in the Offenders Index Codebook (Research Development and Statistics Directorate, 1998). More recent versions of the codebook exist but are not publicly available.

1.3 Motivation for a new measure

While OGRS has been extremely successful, there is a need to develop a new measure. The reasons for this are as follows:

- *The need for a reoffending score.* Offending behaviour rather than convictions is the primary focus of interest. Hence, one needs a measure that more closely reflects this primary focus.

- *The existing score is somewhat outdated*. The measure was constructed in 1998, and the nature of crime and society will have changed substantially in this period. Scores such as OGRS need periodic updating.

- *Problems with the existing OGRS score*. The current OGRS score is thought not to produce accurate predictions for juvenile offenders, and it is reported that the score overpredicts those sentenced to longer custodial sentences (Home Office, 2004; Rose [TO ADD]). Similar verbal concerns have been raised relating to female offenders and for older offenders.

- *Problems with the Offenders Index*. The existing OI database can be criticised in a number of ways:
  - It does not contain true criminal histories. The court-based recording system is simply a record of names and convictions. Criminal histories are formed later by record linkage, through matching on surname, initial and date of birth. This can lead to inaccuracies in the resulting criminal records, particularly those with common surnames and for females (Francis and Crosland, 2002).

  - The system is court-based and does not record cautions, warnings, etc. Thus offences which have been admitted to through police caution and other police disposals (and thus effectively proved) do not form part of the Offenders Index.

  - Only standard list convictions are recorded. Thus more minor offences do not normally appear on the OI, except when they co-occur with a standard list offence.

  Thus, for juvenile offenders, in particular, the OI is perceived as lacking crucial information on potential predictors.

- *Lack of documentation.* For such a widely used score, the lack of documentation on the research methodology and the detail of its construction is somewhat surprising. For the current score (OGRS2) the report by Taylor (1998) presents an overview of the score, but not on the construction of the score.

- *Difficult to use by practitioners.* This issue has been raised by Stephens and Brown (2001) who identified many instances where practitioners, when constructing an OGRS score for an individual, miscoded the criminal history of an individual (for example, by including non-standard list offences). Partially, this is due to the lack of documentation mentioned above (for example which offences belong in which of the 27 offence groups) but also there may be difficulties with some OGRS concepts, such as the definition of principal offence. This whole area is less easily dealt with, but good definitions and examples will play a part in improving accuracy.

1.4 Aims of the current project

This project is concerned with developing a new reoffending score, based on information from the Police National Computer (PNC). As part of this development, the project will aim to improve the documentation of the score, and to make it more transparent. The project will also aim to improve the prediction of the score for subgroups of offenders who are perceived as being poorly predicted by the current model.

The next section (Section 2) outlines the PNC database, and raises some issues with the database which will need to be confronted. Section 3 addresses these concerns. Section 4 presents some initial analysis, and section 5 presented the development of a new reoffending score.

2. Issues relating to the use of the PNC data

The Police National Computer data offers a better coverage of contemporary criminal histories and appears to deal with many of the disadvantages of the OI. For example, among the advantages are: cautions and warnings are recorded, the database covers Scottish offending as well as England and Wales, and criminal histories are drawn from operational records and will be more reliable than those from the OI. Table 1 below, taken from Francis and Crosland (2002), outlines some of the advantages and disadvantages of the PNC.

Many of the disadvantages outlined in the table below have since been addressed:

a) *Disposal codes and offence codes differ*. A comprehensive code conversion scheme within the Home Office has reduced this problem. The ACPO offence codes and disposal codes have [HOME OFFICE to ADD].
b) *Direct searching of PNC database.* Direct searching allows the criminal histories of subsets of offenders who have particular characteristics (such as young, persistent offenders, or those cautioned or convicted of racially aggravated assault) to be obtained directly from the PNC. Procedures are now in place to allow more efficient data extraction for individual records, and it is now possible to extract cases which conform to certain characteristics.
c) *Delays in data entry*. [WHAT CAN THE HO SAY ABOUT THIS?]

This section is therefore primarily concerned with issues which will affect the new measure. These are a) the weeding policy of criminal convictions, b) the more detailed information on offences, c) the relative lack of historical information through back-record conversion and d) the inclusion of Scottish offences and offences from other police forces.

Table 1 Advantages and Disadvantages of the PNC (from Francis and Crosland, 2002)

Advantages	Disadvantages
Includes Scotland, N. Ireland.	Cannot search the database directly
Complete history for older offenders if back records are converted	Offence codes and disposal codes differ from standard RDS codes
Police-based information usually available more promptly.	Delays in data entry on convictions.
Criminal histories built up by fingerprint verification	Criminal histories might be split.
Cautions, warnings and impending prosecutions available	Criminal histories weeded – less important offences removed
Dates of offence available	Disposal information weaker
All offences recorded	Criminal histories deleted on death

a) *The PNC weeding policy*. Weeding is the partial or total removal at some point in the past of a criminal history from the Police National Computer. Convictions become spent after a certain period of time because of the need to rehabilitate offenders; this is a specific requirement under the Rehabilitation of Offenders Act 1974. However, spent convictions will usually stay on the Police National Computer. Weeding is therefore procedural – records are removed following ACPO guidance to reduce the size of the database. The Offenders Index does not suffer from weeding problems, as it is a research database derived from court records, and is private to the Home Office. However, the PNC, as an operational database, will be weeded. This is potentially a serious problem, as it means that information stored on the PNC will not accurately reflect the true criminal history.

We describe the current weeding policy in Table 2 below as determined by the Association of Chief Police Officers (ACPO) and reported in Disclosure News (Criminal Records Bureau, 2003). The main concern is with the loss of historical information. Specifically, on the PNC, information on the date of first caution or conviction will, for many offenders, be lost. In addition, many other OI variables, such as breach history, will become more difficult to implement. We return to this issue later in Section 3, but there are in effect two choices – to summarise the recent criminal history (say within 5 years of a conviction), or to attempt to use the full criminal history, recognising the limitations. It should be recognised that if the latter course is followed, then many variables will change their meaning (thus date of 1st conviction will become date of earliest conviction recorded on PNC).

Table 2. The current weeding rules for PNC data.

Under the weeding rules, where a person has not been convicted of a recordable offence in the past 10 years, his or her record will be deleted unless any of the following conditions applies:

1. *The record contains a total of six months or more imprisonment, including suspended sentences. The total will be the aggregate of all sentences, irrespective of whether they are consecutive or concurrent.*
2. *The record contains three or more convictions for recordable offences.*
3. *The subject has on any occasion been found unfit to plead by reason of insanity or has been sentenced under the Mental Health Acts.*
4. *The record contains a conviction for offences involving indecency, sexual offences or violence, or trafficking in, importation of, or supply of all classes of drugs, or possession of class 'A' drugs.*
5. *The record contains a conviction for an offence involving, as a victim, a child or young person, or one who is elderly, or who is mentally or physically disabled, where the modus operandi indicates that the offender deliberately targets this class of victim.*
6. *The record contains a conviction for an offence involving terrorism under any provisions of anti-terrorism legislation.*

Where condition 2 applies, the record will be kept for 20 years from the date of the last conviction. Where any of the other conditions apply, the record will be retained until the death of the subject, or until the subject reaches 100 years old.
Cautions, police reprimands and final warnings also form part of the police record. If there are cautions but no convictions on the record and no further cautions have been recorded for a period of five years, the record will be deleted, except where the caution is accompanied by an 'offence against vulnerable person' information marker. If there are police reprimands or final warnings but no convictions on the record, the reprimands and warnings will be retained until the offender has attained the age of eighteen years and for a minimum period of five years. After attaining the age of eighteen years and if no police reprimands or final warnings have been recorded for a period of five years, the record will be deleted.

b) *More detailed offence information*. Whereas the Offenders Index contains details on convictions for standard list offences only, the PNC includes information on cautions and convictions for all offences. While this increase in information might be considered to be an advantage, we need to be aware of the implications of this. Specifically, there are two implications:
   - The need to provide a categorisation of all offences into offence groups. The current OGRS formula makes use of 27 specific offence groups which classify the principal target offence. Users of the formula will attempt to classify principal offences which are non-standard list into this categorisation, and we therefore need to calibrate our model using an enhanced categorisation.
   - The inclusion of minor motoring offences and offences such as prostitution might well distort the prediction formula.

c) *Lack of historical information.* The process of back record conversion is imperfect, and information currently recorded routinely on the PNC may not be available on paper records. First among these is the start and end date of offence. While this information is present in around 92% of PNC records after 1996, it is present on less than 2% of records prior to 1996. Using date of offence as a replacement for date of conviction needs to be carefully considered.

d) *Inclusion of Scottish offences and offences from other police forces.* Again, the availability of Scottish offences might be seen to make a criminal history more complete. Those offenders who offend both north and south of the border potentially now have a complete PNC criminal history rather than a partial OI history. In the past, there have been unresolved problems, such as the lack of a conversion scheme between Scottish offences and England and Wales offences, and between Scottish and England & Wales disposals. Scottish offences, for example, have their own set of codes on the PNC, as they relate to Scottish law, and not England and Wales law. However, conversion schemes developed by the Home Office are now in place, and we are therefore able to use complete criminal histories covering both Scotland and England and Wales.

Having identified the main problems with the PNC, we now proceed in Section 3 to operationalise the variables needed for the construction of the new measure.

3. Developing a new measure

It is at this point that we need to confront a fundamental issue. Is the new measure simply a development of OGRS - and, hence, could be easily named as OGRS3 - or is the new measure a rather different beast that, for clarity, needs to be given a different name. While it would be tempting to follow the rigorous pedigree of OGRS, we strongly argue that the latter approach is the more appropriate route.

The proposal of a newly-named measure comes about largely for two reasons. Firstly, there is a genuine re-conceptualisation of the approach to measuring - that is, moving from conviction data to a greater focus on offending behaviour. Secondly, a fundamental re-naming removes the danger of spurious attempts to compare the outcomes of the two measures. While we must have regard to whether or not the new measure is in some senses an improvement on the old measure, we would argue that we cannot compare an apple to an orange. Thus, we have not produced a rather better apple but a rather different fruit.

To take us on our journey of developing a new measure, we need to confront various tasks. First, we define some terminology. Secondly, we need to specify precisely what is meant by reoffending and, thirdly, we need to operationalise what predictor variables can be used. Finally, we need to consider what to name the new measure. We discuss each of these problems in turn.

3.1 Terminology used.

The OGRS score is used to determine risk of offender reconviction at a particular point in time for offenders. Normally this will be at the point of sentence or caution, warning or reprimand. However, for those offenders who receive custodial sentences, this point will be at the release date. We refer to the date of this event as the **sanction/release date**. The age of the offender at this date is referred to as the **current age**. We refer to the **target sanction** or **target conviction** as being the set of one or more offences at the most recent caution or conviction. The **target principal offence** is the principal or most important offence at the target sanction defined using Home Office rules (see for example Criminal Statistics, 2003, p98). **The target sentence** is that sentence which is given to the target principal offence.

3.2 How is "reoffending" assessed?

The focus of the study is to consider two-year reoffending and one-year reoffending as outcome measures. However, the availability of PNC data which contain dates of offence gives a number of choices, and date of reconviction (which was used in previous OGRS measures) may not be the best choice.

The major changes are:

- *Date of offence.* Start date of offence and end date of offence (for offences which are "continuous" offences occurring over a period of time rather than distinct events), which were previously unavailable on the Offenders Index, are now both available on the PNC. Generally, the end date of offence is identical to start date of offence. Date of offence is well recorded for all offences after 1996. The crucial issue to recognise is that, in general, using date of offence is preferable to using date of conviction because it relates more closely to offending behaviour. Moreover, Copas and Marshall (1998) make the point that pre-sentence reports refer to the risk of reoffending and not reconviction; but their data source had no information on dates of offending. In other words, they recognise *our* aim as the essence of the task, but there is a serous shortcoming in the data source of the OI in being able to deliver on the underlying aim. The PNC largely overcomes this limitation.

- *The conviction events.* Using PNC data, the term 'conviction' is unnecessarily restrictive, for there is scope also to consider cautions, warnings and reprimands. By broadening the scope of the disposals, we get much nearer to the notion of offending behaviour. We use the term “sanction” to refer to cautions, warnings and reprimands as well as convictions.

Our choices for “date of reconviction” are therefore: (start) date of reoffence, or date of resanction. Using date of offence is preferable as (because of the time gap between offending and a court appearance) it will allow more offending to be included in any follow-up period. However, moving away from a focus on convictions and attempting to get closer to offending behaviour has its dangers. Indeed, we need to be aware that the PNC is an operational database, and therefore will contain offences which have yet to be proved, either in a court of law or by the suspect admitting guilt through a caution, warning or reprimand. We therefore recommend that “date of *proven* reoffending” is used. We consider all offending which occurs within the two-year or one-year follow-up period, but insist that this offending is “proved” either by a caution (or other police action) or by a court conviction. This means that there needs to be an additional period (which we term a “confirmation period”) in which offences which occur towards the end of the two-year (or one-year) follow-up period have a chance of being proved and appearing on the PNC database.

It should be noted that there are two conditions which need to be satisfied for any offence to be proved –the offender needs to be found guilty (or admit to the offence through caution, warning or reprimand) and this information needs to appear on the PNC database.

Figure 1 illustrates the problem. Case 8 is straightforward – an offence occurs, a court conviction occurs later, and the information appears on both the OI and PNC at different times. This case would be treated as a reconviction in both the OI and the PNC analysis, using either measure of reconviction. However, case 2 is different. The offence occurs early on in the follow-up period, but is brought to court only after the two-year follow-up period is over. However, the court conviction is within the confirmation period and the offence also appears on the PNC database within this period. This would be accepted as a reoffence with the new definition, but would not be a reconviction under the old OGRS definition. Finally, case 7 illustrates that some “valid” reoffending will still fail to be detected – the court conviction for this case occurs within the confirmation period, but fails to be entered onto either the PNC or the OI databases.

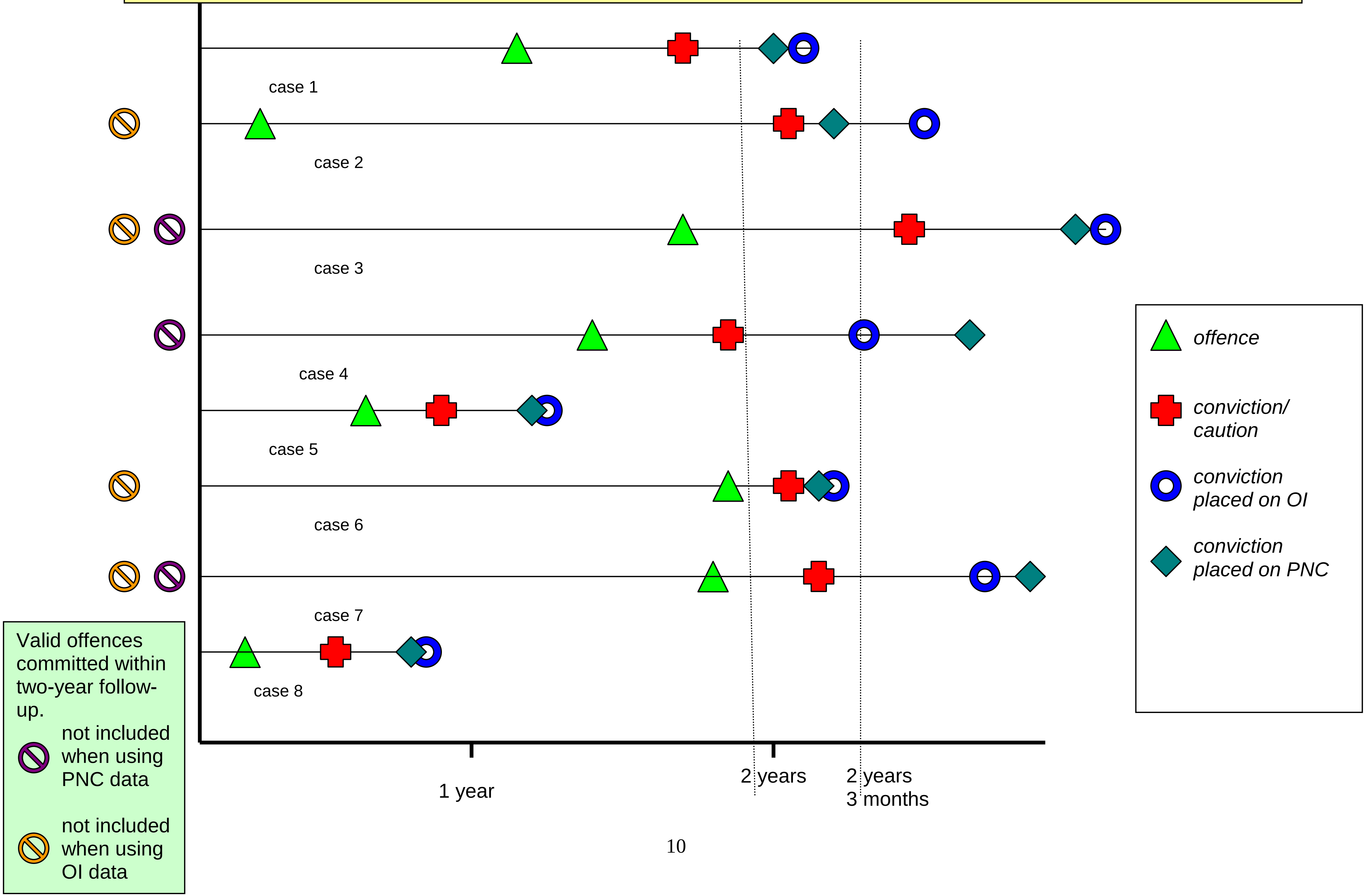
It is therefore important to define both the follow-up period and the confirmation period. The longer the confirmation period is, then the greater the proportion of offenders who will be proved in court.

For this study we therefore use the following as a definition of reconviction/reoffending.

> *An offender who has committed a recordable offence within the follow-up period and who has had the offence "proved" within the follow-up period and a confirmation period of X months, either by the offender accepting a caution, warning or reprimand, or by being found guilty in a court of law.*

We address the issue of the length of the follow-up in section 4.2.

Figure 1. Two year follow-up with three month confirmation period – valid cases included and excluded when using OI and PNC



3.3 What predictor variables can be used?

We now turn our attention to the set of predictor variables. We consider a large set of potential candidate variables – either variables which have been used as predictors in OGRS studies, either by the Home Office or by the Northern Ireland Office, or potential new variables suggested by this research. For each variable, we discuss whether each of the variables still provides a sensible predictor variable in the context of moving to the PNC database, and if so, how to operationalise the variable..

3.3.1 Predictor variables used in the previous OGRS score

We refer to Taylor (1999), which lists the variables used in the current OGRS score. In fact, there are two OGRS scores documented in this report. The first of these is a general score which is usually referred to as OGRS2. The second OGRS score, which is appropriate for sexual and violent offenders, is little used. We focus on the first OGRS score but consider predictor variables from both scores.

The predictor variables used in this analysis are shown in the headings below, and within each section we discuss changes from the original variables used by Taylor (1999).

a) ***Current age***

Taylor (1999) referred to the offender's age in years at time of sentence. Age can now be measured either at the time of offence, at the time of baseline sanction or at the sanction/release date. Initial analysis of the datasets (section 4) has shown that date of offence is not recorded on the dataset for offences prior to 1996. Moreover, it is current age, or age at sanction//release which should determine the risk of subsequent offending – for custodial offenders the age at release is more relevant for risk than the age at sentence. Thus, we take the **current age – the age at sanction for non-custodial offenders and age at release for custodial offenders** – as being the best indicator for this variable.

b) Gender

This can be used straightforwardly and is unchanged from Taylor (1999).

c) ***Current offence group***

The present OGRS2 coding of current offence group proceeds by first determining the principal target offence and then categorising this offence into one of 27 offence groups. The principal offence is defined by the Home Office in *Criminal Statistics England and Wales 2003* (Home Office, 2004) as follows:

> (a) where a defendant is found guilty of one offence and acquitted of another, the offence selected is the one for which he is found guilty;
> (b) where a defendant is found guilty of two or more offences, the offence selected is the one for which the heaviest sentence is imposed;
> (c) where the same disposal is imposed for two or more offences, the offence selected is the one for which the statutory maximum penalty is the most severe.

There are three issues to confront. The first issue is that the previous OGRS score was concerned only with *standard list* offences. With the PNC, we now need to consider all offences, whether standard list or not. While we could place all non-standard list offences into a separate category of "summary non-standard list offences", it has been reported by Stephens and Brown (2001) that, for practitioners and probation officers, determining whether or not an offence is "standard list" is a major source of error in implementing OGRS from paper records.

The second issue is how to determine the principal offence. The PNC extracts provided contain an indicator variable labelled "Primary (Offence)" - this variable is meant to determine principal offence, and the accuracy of this flag is considered later in this report.

The third issue is the definition of the offence categories. We have taken the opportunity to reduce the number of categories from 27 to 19, improving the assignment of offences to categories, removing hrad to define categories and introducing a new category for drink-driving.

We have therefore placed all offence codes into the new 19 offence groups. The list of Home Office offence codes associated with each offence group can be found in Appendix B.

d) ***Age at first recorded sanction***

Taylor (1999) considered " Age at first conviction". With PNC data, there is a choice for this variable on whether to use age at first sanction or age at first proven offence. While using age at first proven offence may be more criminologically correct, the issue here is that we may have lost offences and sanctions because of weeding. Moreover, there is little information on date of offence before 1996 (see Section 4), so using age at first offence would be problematic for older offenders. We have therefore operationalised this variable as age at first recorded sanction.

e) Any recorded prior burglary sanctions

The weeding of criminal records may cause a problem for this original variable of "history of burglary", as early burglary sanctions may have been weeded out. While we considered using "Any burglary convictions in the last five years" as an alternative, we operationalised the variable as "any recorded prior burglary sanctions".

f) Any recorded prior breach sanction*s*

Weeding is also an issue with the variable "history of breach"; furthermore, there may be additional offences included as the analysis moves from standard list offences to all offences. We operationalised this as "any recorded prior breach sanctions".

g) Number of prior violent offences or sanctioning occasions

For the variable "number of prior violent convictions", as well as taking account of weeding problems and the need to integrate new non-standard list offences, we can also move from using convictions to using offences. We operationalised the variable as "number of prior violence offences".

h). Number of prior sexual offences or sanctioning occasions

The issues for the "number of prior sexual convictions" variable are similar to those identified under g). We now look at prior offences rather than convictions, and therefore operationalised this variable as "number of prior sexual offences".

i) The number of offences or sanctioning occasions leading to custody for those under 18.

Four variables were initially considered when operationalising the original variable of "number of youth custodial sentences". Two of these variables relate to *general* custodial sentences given to the offender when under 18. the first of these counts the number of *sanctioning occasions* that resulted in a custodial sentence, and the second counts the number of *offences* leading to custody. The other two variables relate to custody in a Young Offenders Institution (or equivalent forms of youth custody). The first measures the number of *sanctioning occasions* that resulted in a sentence in a YOI. The other counts the number of *offences* leading to custody in a Young Offenders Institution. We used the first two of these "the number of offences leading to custody for those under 18" and "the number of sanctioning occasions leading to custody for those under 18".

j) Number of offences (NO) and sanctioning occasions (NS)

The variable "number of convictions" is used in the calculation of the Copas rate. With the move to the PNC, we can choose to look at either offences or sanctioning occasions – we have chosen to count offences as the more relevant measure.. Weeding policy may mean that older data is missing. We operationalised this variable in two ways – either using "Number of offences (NO)" –this includes the offences at the target sanction – or number of sanctioning occasions, including the target sanction

k). Length of criminal history in years from first conviction (T)

For this variable, we have kept to the original definition. Replacing date of first conviction by date of first offence, was considered, but would be too unreliable as date of offence is often missing. Hence, we operationalise as "length of criminal history in years from 1st conviction to current date".

l) The Copas rate.

The Copas rate has been used on all previous OGRS scores, and is defined to be a transformation of the prior offending rate. If we define *NC* to be the number of prior convictions (in our terminology, sentencing occasions) committed up to the assessment time, and *T* to be the length of the criminal history, then the Copas rate was defined to be

$$\sqrt{\frac{NC}{T + k}}$$

where *k* is an "offset" factor which applies a correction to the length of criminal career. This offset factor can be thought of as a correction to the observed criminal career length as the true length of criminal career is not known. For previous versions of OGRS, *k* has been set to five years.

We can see that the square root transformation works to discount the effect of large numbers of convictions on the risk. Thus, the effect of an *extra* conviction on the risk of resanction is less for someone with 100 previous convictions compared to someone with 10 reconvictions.

Use of the PNC data will mean that the Copas rate could use either the number of offences NO or the number of sanctions NS, and this in turn may mean that a different constant *k* and a different transformation may be needed. This will be investigated as part of the analysis. Assuming a square root transformation for the time being, the Copas rate used in this analysis will be either

$$\sqrt{\frac{NO}{T+k}} \text{ or } \sqrt{\frac{NS}{T+k}}$$

3.3.2. Variables used in the Northern Ireland OGRS measure.

Recent work by Francis, Harman and Humphreys (2005) has developed an OGRS measure for use in Northern Ireland. The aim of the Northern Ireland work was somewhat different, and was focused more towards the development of a forecasting tool to predict future reconviction so that the Northern Ireland Public Service Agreement on Criminal Justice (which aims to reduce reconviction by 5% by April 2008) could be implemented.

Mostly, variables were used in that study which have equivalences in the variable list above. One crucial difference was the development of a separate model for non-custodial and custodial offenders. Additional variables which were considered fell into three categories:

a) Those which relate specifically to Northern Ireland (Scheduled/non-Scheduled offence)
b) Those relating to type of disposal.
c) Those relating to length of custodial sentence.

We considered that length of custodial sentence might well prove to be an important variable given the perceived lack of fit for those serving long custodial sentences. However, we did not wish to include a variable which discriminated in fine detail between custodial sentences. We therefore included the following variable:

m) Whether the principal offence of the target conviction had a custodial sentence of 4 years or more.

3.3.3. Additional variables considered.

n) Number of cautions, warnings and reprimands in the criminal history.
Given the perceived lack of fit for young offenders, we decided to include two alternative additional measures which assess the history of the offender at the target conviction in greater detail. We considered two such measures:

1. the number of cautions, warnings and reprimands in the criminal history. Cautions, warnings and reprimands are subject to weeding, and the accuracy of this variable for some older offenders (whose early history may have been weeded) is therefore suspect.

2. A variable describing the nature of the target sanction and the prior history. We identify whether the target sanction is a first sanction occasion (caution or conviction), a second caution or another type of offender.

3.4 Score development strategy.

Our aim in developing a new OGRS score was to address the criticisms of the previous OGRS scores. We therefore aimed:

a) to improve the ease of use of the score for Probation Officers when using printed criminal histories.
b) to improve predictions for young offenders, custodial offenders and other groups.
c) to use modern analysis methods to ensure rigour.

3.5 Naming the new measure

The earlier discussion in this section identifies that the pedigree of the OGRS approach and the development of a new measure are very closely related. However, there is a basic reconceptualisation that provides the rationale for a renaming process. The OGRS approach - while acknowledging that it is trying to predict the risk of reoffending - is based on conviction data. This new approach is attempting to move away from conceptualising the issue in terms of reconviction but in terms of reoffending. In reality, of course, we cannot know of all the reoffending that takes place which will include illegal activity that is not detected. Furthermore, the currency with which we deal must go beyond suspicious behaviour known to the police, but must focus on *proven* reoffending. Hence, the transition from focusing on conviction to reoffending behaviour will not be complete, but the transition does undermine attempts to compare the performance of the two measures directly.

The name OGRS - the Offenders Group Reconviction Score - highlights the notion of 'reconviction' as central to the endeavour. Our proposed name for the new approach - ASPRO (<u>A</u> <u>S</u>core for <u>P</u>roven <u>R</u>eoffending of <u>O</u>ffenders) - reflects the increased focus on 'reoffending' as the pivotal concept.

4. The datasets

We took two extracts from the Police National Computer for this study. The datasets were downloaded on the 14th February 2005 and consisted of the following samples:

a) A set of 81744 [1] offenders with non-custodial sanctions in England and Wales in January 2002.
b) A set of 7921 offenders released from custody in January, February or March 2002.

As date of release from custody is not recorded on the Police National Computer, the second dataset was constructed by obtaining a list of offenders released from custody, and [HOME OFFICE TO ADD].

For the non-custodial dataset, there were 1536374 separate offences. These offences were committed by 81744 offenders on 697576 occasions. For the custodial dataset, the initial dataset contained 295644 offences. These offences were committed by 7921 offenders on 109174 occasions.

Of all sanctions received on the target offence date, 84.2% were court convictions, 9.3% were cautions, 4.0% were reprimands, and 2.4% were warnings.

Of all sanctions received on the target offence date that were identified as the primary offence, 76.6% were convictions, 14.2% cautions, 6.0% reprimands and 3.2% warnings. The lower rate of court convictions for primary offences perhaps seems puzzling, but can be explained as early sanctions consist of few offences and are usually cautions etc; later sanctions have more offences and will tend to be court convictions.

4.1 Initial analysis of the data

An initial analysis was undertaken which investigated the datasets for consistency. We examined the ages of offenders, the location of the convictions and issues relating to data completeness.

We noted first of all that the dataset had missing dates of birth for some offenders, and for other offenders the dates of birth gave obviously wrong ages of conviction. Missing dates of birth were recorded as 1/1/1900, and we removed all such offenders from the database. Francis, Soothill, Humphreys and Bezzina (2005) noted in a parallel analysis of PNC data that many of these missing dates of birth would be where businesses or companies have been convicted. There were a small number of offenders with age at the target conviction falling between 8 and 10 years (10 years is the age of criminal responsibility in England and Wales). These were offenders who were sanctioned in Scotland, where the age of criminal responsibility is 8 years. These offenders were also removed, together with offenders with target ages of less than 8 years. This removed a total of 51 offenders from the database (48 from the non-custodial dataset and 3 from the custodial dataset).

[1] There were originally 82293 offenders, however, 549 did not receive a conviction within the target period but somehow found their way on to the dataset

We then noted that the dataset supplied contained target convictions for all regions in the UK. As the primary purpose of this analysis was to develop a reconviction/reoffending score for offenders in England and Wales, we removed all offenders whose *target* convictions were in Scotland and Northern Ireland or with a non-Home Office force (such as the British Transport police). However, as discussed earlier, we retained non-Home Office force offences in the *previous history* of the offender. This removed an additional 7169 offenders from the database (7100 from the non-custodial dataset and 69 from the custodial dataset). Of the 7100 offenders in the non-custodial dataset whose target convictions were not in a Home Office force the majority were in Scotland, particularly the Strathclyde force. Of the 69 offenders in the custodial dataset whose target convictions were not in a Home Office force the majority were listed as being convicted by the British Transport Police.

Finally, there were 182 cases with missing gender (all from the non-custodial dataset). Of the remaining cases 3069 principal offence code were missing (2895 from the non-custodial dataset and 174 from the custodial dataset). These were also removed. The final dataset gave a total of 7675 custodial offenders and 71519 non-custodial offenders.

Table 3 shows summary statistics for this final custodial and non-custodial dataset.

Table 3. Summary statistics for the final datasets used for the construction of the new OGRS measure.

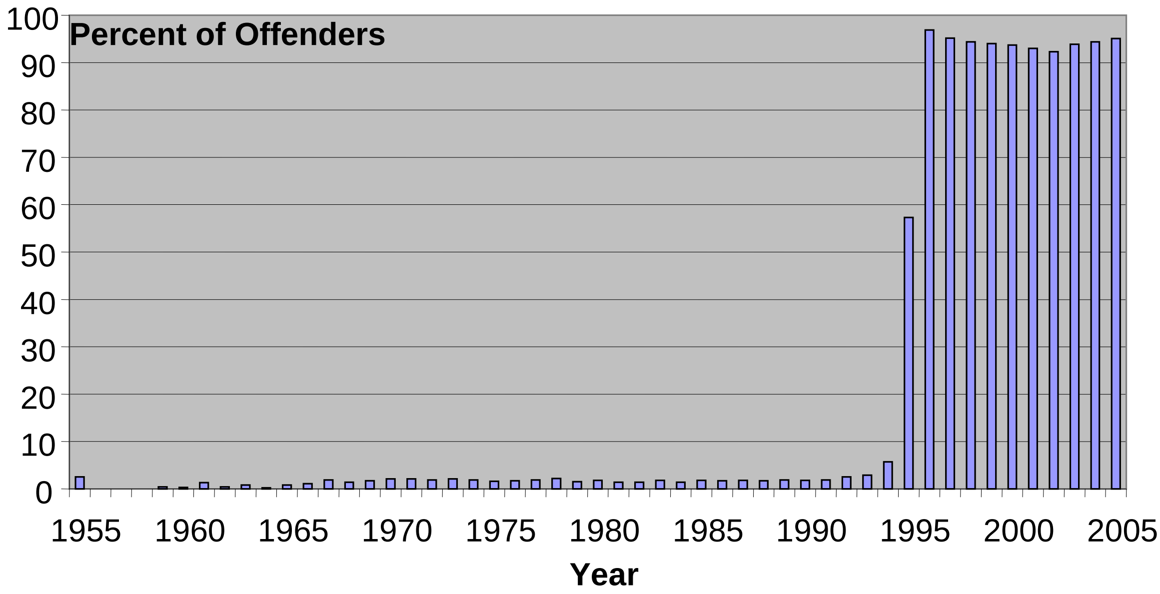
	Non-custodial dataset	Custodial dataset
Average at sanction/release	27.0	29.4
Average number of previous offences	12.0	26.6
Average number of previous sanction occasions.	5.2	9.7
Proportion of males	83.0%	84.0%
Number of offenders	71519	7675

We then carried out an initial analysis to determine the last date of recording. We examined the dataset to investigate whether caution and conviction dates existed for dates close to the download date. All police forces in England and Wales had caution and conviction dates in February 2005. However a small number of forces had substantially lower numbers of cautions and convictions in the later months of 2004 suggesting incomplete entry of information. Nevertheless, despite a few problems, it appeared to us that information on convictions was being added to the PNC in a very timely manner.

A final task was to examine the historical completeness of the data. We noted one feature of the PNC data – that the start date of offence appeared to be unavailable before 1996. Figure 2 shows the percentage of convictions for each year for which information exists relating to the start date of offence. Figure 2 confirms this initial view – less than 2% of offences sanctioned before 1996 had this information. While recording of this information was also not complete, over 92% of offences had this information from 1997. It is worth noting that 1996 was the date at which the paper records of old Criminal Records Office were replaced by the Police National Computer.

Because this date of computerisation is important, we should also examine completeness of the PNC records. For those histories starting before 1996, data needs to be back-converted from paper records, before being placed on the PNC. To check if this process had been carried out, we looked at the caution/conviction dates. A lack of back-record conversion would manifest itself by a sudden drop in offences prior to 1996. No such effect was found, leading us to conclude that the process of record conversion was operating satisfactorily.

Figure 2. Percent of offenders with 'start date of offending present' by year of caution/conviction



4.2 Choice of confirmation period.

We now turn our attention to the choice of confirmation period – that is, the period at the end of the follow-up period which is used to confirm offences as proven. We took all offenders who had reoffended within two years and proven by the download date of 14th February, 2005. For these offenders, we then determined what percentage of offenders were confirmed within certain lengths of confirmation periods. We then repeated this analysis for reoffending within one-year.

Table 4 shows the results, separately for the custodial and non-custodial datasets. In general, the custodial dataset has the higher percentages, suggesting that custodial offenders who reoffend will reoffend quickly. In general, a confirmation period of three months appears to give the most appropriate balance between achieving a high percentage of offenders confirmed, and the need to obtain data quickly. For both custodial and non-custodial data, the percentages are lower for one-year reoffending than for two-year reoffending; this is explained by the fact that a greater proportion of offences occur toward the end of the follow-up period for the one-year reoffending.

We took a confirmation period of three months for both one-year reoffending and two-year reoffending.

Table 4. Percentage of offenders reoffending within the follow-period who are confirmed, for various length of confirmation period.

	1 year reoffending		2 year reoffending	
	Non-custodial	Custodial	Non-custodial	custodial
1 month	87.3%	89.5%	95.9%	96.9%
2 months	90.1%	91.9%	96.9%	97.6%
3 months	92.1%	93.7%	97.6%	98.1%
4 months	93.6%	95.1%	98.1%	98.5%
5 months	94.7%	96.1%	98.5%	98.7%
6 months	95.6%	96.9%	98.9%	98.9%

4.3 Accuracy of the Primary flag on the PNC

Computer downloads of the PNC have a “Primary offence” indicator to indicate which of a set of offences dealt with at the same sanctioning date is to be taken as the principal offence. We investigated this flag to determine its accuracy. Of the sanctioning occasions in the sample, 93% had one primary indicator, but 7% had more than one indicator. The largest number of primary flags for a sanctioning occasion was 80.

We also examined 100 sanctioning occasions in more detail. Most of these had a single primary flag, and we determined that the correct offence had indeed been flagged in each case. However, there was one sanctioning occasion where there were two flags – one for an offence of actual bodily harm and one for an offence of common assault. Clearly, the common assault offence should not have been flagged.

5. Constructing the reoffending measure.

We adopted the following procedure in constructing the new reoffending measure.

We first divided the dataset into two. 60% of the sample (the calibration dataset) was used to construct the reoffending score, and the remaining 40% of the data (the validation dataset) was not used in the model fitting but reserved to determine the accuracy of the score. The *calibration* dataset consisted of 42913 non-custodial offenders and 4602 custodial offenders, making 47515 offenders in total. The validation set consisted of 31679 offenders –28606 non-custodial and 3073 custodial offenders.

Models were then fitted to the calibration dataset. Once a final model had been determined, the model was assessed by comparing the fitted probabilities from the model with the actual reoffending measures for the validation sample

5.1 Logistic regression method – separate models for one-year and two-year reoffending

Logistic regression provides a standard method for modelling the probability of reoffending at a specified time point. We will be able to develop separate models which model the probability of two-year reoffending and one-year reoffending. Logistic regression models the probability of an event (such as two-year reoffending) as a function of a linear combination of selected explanatory variables. It was used in the construction of the original OGRS score and also the more recent OGRS2 score.

a) For the combined calibration dataset, we fitted a series of logistic regression models, modelling the probability of two-year reoffending as a function of explanatory variables. We repeated the process for one-year reoffending.

b) We fitted a wide variety of models and compared them using two criteria.

    1. The Akaike Information Criterion (AIC) (see e.g. Lindsey and Jones, 1998). The AIC is defined to be the value of minus twice the log-likelihood of the fitted model penalised by adding on twice the number of parameters in the model. It aims to strike a balance between improving the fit of a model and having a small number of parameters. Models with a smaller AIC are preferred to models with a larger AIC.
    2. A measure of concordance (sometimes referred to as *c)* which assesses predictive discrimination. This concordance index measures correlation between the predicted probability from the model and the observed outcome. The measure can be calculated by examining all possible pairs of offenders, where one of the pair is a reoffender and one not, and calculates the proportion of these pairs where the predicted probability of reoffending is greater for the reoffender than for the non-reoffender. It has been shown to be equivalent to a measure known as AUC_{ROC} - the area under the ROC curve, which can be

calculated simply in SPSS and other packages. (Harrell, Lee and Mark, 1996; Zhou, Obuchowski and McClish, 2002).

We also needed to consider collinearity, which can cause problems in estimation. Typically, inflated variances for the estimated coefficients are produced once a problematic explanatory variable is included in the model, and this can lead to poor predictions for certain combinations of explanatory variables (Hauck and Donner, 1977). In standard linear regression it is caused by high correlation between the explanatory variables (strictly columns of the design matrix), or more generally, where one column of the design matrix can be expressed as a near-linear combination of other columns of the matrix. However, in logistic regression and other models, such ill-conditioning problems can also be caused by maximum likelihood estimation difficulties relating to the parameter estimates. We adopted the technique of assessing collinearity suggested by Lesafffre and Marx (1993).

c) We first fitted statistical models to reoffending which corresponded as closely as possible to the previous OGRS formulae for reconviction. We utilised two alternative sets of variables relating to previous criminal history (prior sexual criminal history, prior violent criminal history, prior youth custody history and Copas rate) – a set based on counting *offences*, and a set based on counting *sanctioning occasions*. We also tried both *continuous* and *categorical* forms for the continuous variables of current age, age at first sanction, prior sexual history, prior violent history and prior youth custody history. For both the one-year and two-year reoffending, the categorical versions both gave substantially smaller values of the AIC measure compared to the continuous measures (Table 5). We tried various forms of categorisation with categories being merged when they showed little or no difference. Table 6 shows the AIC values for two-year and one-year reoffending, for the two sets of criminal history variables. Using the criminal history variable set based on sanctioning occasions, the two-year reoffending, the AIC was reduced from 51104 to 50967, and for one-year reoffending, the AIC measure was reduced from 49660 to 49514. The criminal history variables based on offences had substantially higher AIC values, but also showed reductions in moving from continuous variables to categorical. The AUC values show the same pattern, although the criterion is less sensitive. The final form of the variables used is given in Table 6.

Table 5. AIC values and AUC_{ROC} for initial models using continuous and categorical forms of explanatory variables, and alternative criminal history variable sets.

	One-year reoffending		**Two-year reoffending**	
	Continuous variables	**Categorical variables**	**Continuous variables**	**Categorical variables**
Criminal history variable set based on counting offences	AIC: 50233 AUC: 0.786	AIC: 49973 AUC: 0.789	AIC: 51721 AUC: 0.801	AIC: 51414 AUC: 0.803
Criminal history variable set based on counting sanctioning occasions	AIC: 49660 AUC: 0.791	**AIC: 49514** **AUC: 0.793**	AIC: 51104 AUC: 0.806	**AIC: 50967** **AUC: 0.807**

Table 6. List of variables used in final reoffending model

AGEBAND Current age	10 and under 12 = 1, 12 and under 14 = 2, 14 and under 16 = 3, 16 and under 20 = 4, 21 and under 25 = 5, 25 and under 30 = 6, 30 and under 35 = 7, 35 and under 40 = 8, 40 and under 50 = 9, 50 and over = 10.
FIRSTAGE Age at first sanction	Under 12 = 1, 12 and under 14 = 2, 14 and under 16 = 3, 16 and under 20 = 4, 21 and under 25 = 5, 25 and under 30 = 6, 30 and under 35 = 7, 35 and under 40 = 8, 40 and under 50 = 9, 50 and over = 10
LENG Length of criminal career	0 = 0, 1 = greater than 0 and up to 2 years, 2= greater than 2 years
PRECAUT Previous cautions	0 = 0, 1 =1, 2 and over = 2
PVIOL Number of prior violent sanctions	0 = 0, 1 and over = 1
PSEX Number of prior sexual sanctions	0=0, 1 and over = 1
PREVCUST Number of previous youth custodial sanctions	0=0, 1 to 3 sanctions = 1, 4 or more sanctions=2
ANYBURG Any prior burglary	No = 0, yes = 1
ANYBRC Any prior breach	No = 0, yes = 1
CUST4	No prior custody or prior custody under 4 years = 0, 4 years and over = 1.
TARGOFF	1 = Violence, 2 = Robbery, 3 = Public order/riot, 4 = General sexual offences, 5 = Child sexual offences, 6 = Solicitation/prostitution offences, 7 = Domestic burglary, 8 = Other burglary, 9 = Theft, 10 = Handling, 11 = Fraud and forgery , 12 = Absconding and bail offences, 13 = Taking a vehicle without consent and related, 14 = Theft from car 15 = Other motoring, 16=Drink driving 17 = Criminal damage , 18 = Drugs import/ export /production, 19 = Drugs possession and supply, 20 Other
COPAS	Copas rate
TARGHIST Target sanction history	1= Target sanction is first caution, warning, reprimand with no previous sanction 2 = Target sanction is second caution, warning, reprimand with no earlier convictions. 3 = Target sanction is first conviction with no prior sanctions 4= Target sanction is any other caution, warning or reprimand 5= Target sanction is any other conviction..

d) Based on the analysis in the previous section, it was clear that using the categorical form of the explanatory variables with criminal history variables based on sanctioning occasions offered the best way forward. Within this constraint, we then investigated various measures of including the length of criminal history in the model.

The alternative methods were either to include age at first sanction in the model or to include length of criminal career. As current age was already in the model, it was not desirable to have both age at first sanction and length of criminal career present (as current age is equal to age at first sanction plus length of criminal career). We categorised length into three broad categories (LENG)– zero length, greater than zero and up to two years, and greater than two years. We then replaced age at first sanction with length of criminal career. The results improved the model greatly. For two-year reoffending the AIC decreased from 50967 to 50815 for the two-year reoffending measure, and from 49514 to 49356 for the one-year reoffending model.

Examination of the parameter estimates of length of criminal career appeared to suggest that the most important distinction appeared to be between those with no previous sanctions (with a criminal career length of zero) and those with a previous sanctioning history (with positive length of criminal career). We therefore replaced the variables LENG and previous custody (PREVCUST) with a new five-level categorical variable TARGHIST, which categorised offenders on the nature of the target sanction according to prior history. This produced a further decrease in AIC from 50815 to 50747 for the two-year reoffending measure, and from 49356 to 49271 for the one-year reoffending model.

e) We then turned our attention to the Copas rate. We fitted various forms of the Copas rate, which was based on the total number of sanctions NS (including the target sanction) and the length of criminal career T: $\sqrt{\frac{NS}{T+k}}$. We varied the parameter *k*, allowing it to take values from 1 to 50, and also allowed the square root transformation to be substituted by other transformations – alternative transformations tried were the log transformation and no transformation at all. The results showed that the log transformation was preferred over both the square root transformation and the identity transformation for both the two-year and one-year reoffending models. For the two-year reoffending model, the AIC value was at a minimum at k=11. For the one-year reoffending model, a minimum AIC was found at *k*=9. We therefore recommend the use of a natural log transformation rather than a square root transformation when using PNC offending data. The AIC values are given in Table 7.

f) We also investigated whether the Copas rate could be replaced in the model by a term $\ln(NS)$ - the natural log of the number of sanctions, removing the length component from the Copas rate. Putting this term in the model in place of the Copas rate gave an AIC of 49543 for one-year reoffending and 51089 for the two–year reoffending model. This replacement appears to be unacceptable for both the one-year and two-year reoffending models (as there is a large increase in AIC) and we chose instead to use the Copas rate.

Table 7. AIC values for various forms of Copas rate.

	AIC for 1-year reoffending			AIC for 2 – year reoffending		
Form of Copas rate	Transformation			Transformation		
k	Identity	Square root	Natural log	Identity	Square root	Natural log
1	50130	49958	49906	51772	51617	51591
5	49533	49271	49198	50979	50747	50742
6	49517	49246	49161	50956	50709	50685
7	49510	49234	49141	50946	50688	50650
8	49510	49229	49131	50945	50677	50629
9	49513	49229	**49127**	50948	50674	50617
10	49518	49231	49128	50955	50676	50612
11	49525	49236	49132	50963	50680	**50611**
12	49533	49242	49138	50973	50686	50614
15	49558	49264	49160	51006	50711	50632
20	49601	49301	49200	51060	50756	50673

g) We aim for a common Copas rate variable which can be used for both the one-year and two-year model. It is clear that the natural log transformation is most appropriate for both the one-year and two-year reoffending models. In choosing a value of *k*, we seek a compromise between *k*=9 for the one-year model and *k*=11 for the two-year model. We chose a value of *k*=10.

h) The revised model with *k*=10 was then fitted. This gave AIC values of 50612 for the two-year offending model and 49128 for the one-year offending model.

i) Finally, we carried out a backward elimination procedure on both the one-year and two-year reoffending models. Terms were removed from the model if they reduced the AIC value, with those producing the greatest reduction being removed first. For the one-year reoffending model, the variables PVIOL, ANYBURG and PSEX can all be removed in turn, giving a reduction in AIC from 49128 to 49122. For the two-year reoffending model, the variables PVIOL, ANYBURG and ANYBRC can all be removed in turn, giving a reduction in AIC from 50612 to 50610.

j) However, we notice that it is worth removing the marginally significant variables PSEX and PREVCUST from the two-year reoffending model and ANYBRC from the one-year model to give a common explanatory model for both outcome measures. The AIC values change to 49125 for the one-year model and 50611 for the two-year model. To ensure consistency between the two models we also remove PREVCUST from the one-year model – this increases its AIC to 49141. The final AUC_{ROC} values on the calibration dataset are 0.796 for the one-year model and 0.810 for the two-year model.

The parameter estimates for these main effects models are presented in Appendix A1. We can see a number of features. For two-year reoffending, the estimates for target age band show an increase from 0.0 at the lowest age band, peaking in the current age band 12-14, and then declining. This represents the classic age-crime curve. For one-year offending estimates are similar in the lowest three age groups before declining in a similar way. A target conviction rather than a caution and being male both tend to increase reoffending. The estimates also show that those on their first sanction (caution etc. or conviction) or second caution have a substantially lower risk of reoffending. Finally, those sentenced to a long period of custody of 4 years or greater have a reduced risk of reoffending with all other factors equal.

We notice that the parameter estimates in general are very similar for the one-year and two-year reoffending models (apart from the constant term) and this leads to the possibility of estimating a combined model, with common parameter estimates for the two models and differing only in the constant term. We consider this possibility in the next section.

5.2 Ordinal regression – building a single model for one-year and two-year reoffending.

Ordinal regression provides a method of combining the two models above into a common score, allowing a single score to be used for predicting both one-year and two-year reoffending. The method uses an ordinal response variable *y* – taking the value 1 for recidivism within one year, 2 for recidivism over one year but within two years, and 3 for no recidivism within two years - which provides an ordered sequence of responses. We then model the two cumulative probabilities – the probability of being in category 1 and the probability of being in categories 1 and 2 - in a common model.

One particular form of the ordinal regression model – the proportional odds model with logit link- has strong similarities to logistic regression, and takes the form:

$$\log\left(\frac{prob\ (y \le j)}{1 - prob\ (y \le j)}\right) = \alpha_j + \sum_{k=1}^{K} \beta_k x_k$$

where $x_k, k = 1 \ldots K$ are a set of predictor variables, and β_k and α_j are parameters to be estimated. We are interested in j=1 – the probability of reoffending within one year, and j=2 – the probability of reoffending within two years. The estimated models differ solely by a constant - α_1 for j=1 is replaced by α_2 when j=2.

As an illustration, we can look at the actual number of offenders and their proportion falling into each of the three ordinal groups for males and females (Table 8)

Table 8. Reoffending status for calibration sample, by gender

	Reoffended within one year (j=1)	Reoffended between one year and two years (j=2)	Not reoffended within two years (j=3)
Males	14447 36.5%	5271 13.3%	19850 50.2%
Females	2363 29.7%	729 9.2 %	4855 61.1%

We can see for the males that 36.5% have reoffended within one year, and 49.8% within two years. For the females, 29.7% have reoffended within one year and 38.9% within two years. There appears to be a gender difference, which can be formally tested using the ordinal regression model.

The model can be fitted using standard software, including SPSS. We use the same set of variables as given in appendix A1. In a similar way to the logistic regression analysis, the AIC criterion can be used to determine the best set of variables, but these are not comparable with the AIC values for the logistic regression. However, we can also produce comparable "pseudo AIC" values by calculating the fitted probabilities for reoffending after one year, and using the formula:

$$\text{pseudo AIC}_{ord}(1) = n - 2\sum_i y(1)_i \log(\hat{p}(1)_i) + (1 - y(1)_i).\log(1 - \hat{p}(1)_i)$$

where n is the number of parameters in the model[2], $y(1)_i$ is the observed indicator of whether reoffending for case i took place or not within the one-year follow-up time (1=yes, 0 = no) and $\hat{p}(1)_i$ is the fitted probability of reoffending from the model. A similar formula can be used for two-year reoffending.

We chose the final set of explanatory variables used in the logistic regression (using a Copas score with log transformation and k=10), and carried out a backward elimination procedure, removing the least significant variable at each stage until there was no further decrease in the AIC value. The AIC for the full model is 62436. The parameter estimates suggest that PVIOL and ANYBURG are strong candidates for removal. Removing these two variables gives an AIC of 62432 – a smaller value. Additionally excluding PSEX, ANYBRC and PREVCUST gave an AIC of 62439 – a slightly higher value but not a great change. As already noted these AIC values are not comparable with the logistic regression AIC values presented earlier. Psuedo-AIC values were calculated for this final model and gave a pseudo-AIC of 49152 for one-year reoffending and 50623 for two-year reoffending, which show that the fit does not worsen substantially in moving to the ordinal model from the two individual logistic models.

The final model is given in appendix A2. The parameter estimates are very close to those for the logistic regressions. The model appears simpler to use, and we proceed to the next section where we compare how well the two approaches perform on the validation data.

[2] *n* rather than *2n* is used in this formula as estimated regression parameters are common to both the one-year and two year fitted probabilities.

6. Assessing the accuracy of the models

We compared the models by looking at the *validation* dataset, using the 40% of offenders in the sample which were not used to fit the model. This provides a picture of how the reoffending models perform in practice.

We examined two criteria:

1) The area under the ROC curve AUC_{ROC} for the final fitted models restricted to the 40% validation sample.

2) The fitted percentage of offenders reoffending after one year and after two years (both with a three month confirmation period) with the estimated percentage of offenders predicted by the model, again restricted to the 40% validation sample. We considered three ways of classifying offenders: a) by age band and gender; b) by target offence type; and c) by length of criminal career.

6.1 Comparability of validation dataset and calibration dataset

We first compared the validation and calibration datasets by looking at one-year and two-year reoffending rates by gender and agegroup. The table below gives the observed reoffending rates.

Current age	calibration 60% of data	Validation 40% of data	One-year reoffending		Two-year reoffending	
MALES	N	N	calibration %	validation %	calibration %	validation %
10 and under 12	253	164	0.344	0.250	0.482	0.402
12 and under 14	976	651	0.372	0.369	0.541	0.542
14 and under 16	2426	1636	0.456	0.441	0.616	0.597
16 and under 18	3602	2436	0.476	0.478	0.632	0.620
18 and under 21	6645	4240	0.422	0.436	0.567	0.578
21 and under 25	6580	4421	0.381	0.394	0.527	0.531
25 and under 30	5901	3869	0.381	0.388	0.513	0.523
30 and under 35	4675	3128	0.342	0.347	0.473	0.473
35 and under 40	3385	2329	0.301	0.308	0.413	0.429
40 and under 50	3359	2231	0.235	0.235	0.335	0.324
50 and over	1766	1156	0.120	0.138	0.171	0.189
FEMALES						
10 and under 12	49	32	0.163	0.313	0.286	0.406
12 and under 14	349	239	0.241	0.218	0.344	0.326
14 and under 16	804	525	0.274	0.295	0.378	0.400
16 and under 18	681	501	0.278	0.257	0.367	0.345
18 and under 21	1007	688	0.338	0.337	0.425	0.422
21 and under 25	1182	804	0.372	0.367	0.465	0.461
25 and under 30	1117	774	0.362	0.413	0.473	0.482
30 and under 35	978	668	0.334	0.323	0.424	0.416
35 and under 40	783	491	0.241	0.291	0.330	0.377
40 and under 50	710	489	0.196	0.221	0.276	0.282
50 and over	287	207	0.080	0.063	0.101	0.121
Overall	47515	31679	0.354	0.361	0.480	0.482

We observe first of all that the overall reoffending rates are very similar in the two sub-samples. However, when we examine reoffending rates for individual age-sex combinations, we notice some larger differences. While we have not carried out formal statistical testing, we have identified differences which are larger than four percentage points by shading the cells. We see that the reoffending rates for both one-year and two-year reoffending differ in the youngest age category for both males and females, and in the female 35 and under 40 age group. The one-year reoffending rates for females also show large differences for the 25 and under 30 age group.

While such differences are to be expected in any division of data into two sub-samples (caused either by natural sampling variation combined sometimes with small cell numbers) it is important to consider such differences when interpreting the tables below.

6.2 Area under the ROC curve (AUC) for the competing models.

The area under the ROC curve was calculated for the final models for one year and two-year reoffending, and the results are given below. The results are identical for the two models.

	One-year reoffending	Two-year reoffending
Logistic model	0.793	0.810
Ordinal model	0.792	0.810

6.3 Comparing fitted estimated rates of reoffending with observed.

6.3.1 Reoffending status by age band and gender

Table 9 below gives the actual and predicted reoffending percentages. Actual reoffending rates for males peak in the 16-18 age category before declining slowly as age increases. Reoffending rates for females peak later, in the early twenties. Both the logistic and ordinal models agree well with the actual reoffending rates for both the one-year and two-year analyses, for males. However, both models appear to slightly underestimate reoffending rates for young males and slightly overestimate reoffending rates for young females. Other female categories also seem to be discrepant – notably the 25 to 30 age band. Goodness of fit statistics indicate a certain lack of fit, with the logistic models performing better than the ordinal models.

Table 9. Reoffending status by age-group and gender, comparing actual general reoffending rates with those predicted by the models.

Current age		One-year reoffending			Two-year reoffending		
MALES	N	Actual %	Logistic %	Ordinal %	Actual %	Logistic %	Ordinal %
10 and under 12	164	0.250	0.291	0.285	0.402	0.435	0.433
12 and under 14	651	0.369	0.368	0.375	0.542	0.530	0.522
14 and under 16	1636	0.441	0.423	0.428	0.597	0.578	0.570
16 and under 18	2436	0.478	0.454	0.457	0.620	0.605	0.597
18 and under 21	4240	0.436	0.419	0.420	0.578	0.562	0.555
21 and under 25	4421	0.394	0.385	0.389	0.531	0.529	0.521
25 and under 30	3869	0.388	0.394	0.394	0.523	0.532	0.526
30 and under 35	3128	0.347	0.357	0.356	0.473	0.490	0.485
35 and under 40	2329	0.308	0.304	0.299	0.429	0.422	0.422
40 and under 50	2231	0.235	0.236	0.230	0.324	0.339	0.339
50 and over	1156	0.138	0.118	0.105	0.189	0.166	0.171
FEMALES							
10 and under 12	32	0.313	0.253	0.227	0.406	0.350	0.372
12 and under 14	239	0.218	0.295	0.282	0.326	0.416	0.429
14 and under 16	525	0.295	0.329	0.314	0.400	0.449	0.461
16 and under 18	501	0.257	0.380	0.365	0.345	0.496	0.504
18 and under 21	688	0.337	0.355	0.338	0.422	0.457	0.465
21 and under 25	804	0.367	0.329	0.316	0.461	0.427	0.434
25 and under 30	774	0.413	0.347	0.331	0.482	0.437	0.446
30 and under 35	668	0.323	0.275	0.258	0.416	0.358	0.368
35 and under 40	491	0.291	0.245	0.227	0.377	0.313	0.327
40 and under 50	489	0.221	0.194	0.178	0.282	0.257	0.268
50 and over	207	0.063	0.077	0.061	0.121	0.100	0.110
Goodness of fit X2			67.55 on 20 df	88.92 on 20 df		65.54 on 20 df	65.17 on 20 df

6.3.2 Reoffending status by target offence type

Table 10 gives reoffending status for one and two-year reoffending classified by the 27 types of target offence. This is a new and useful method of examining the fit of various models and the actual reoffending rates. As one would expect, property offences have high rates of reoffending. However, other activities such as absconding from bail, which also have remarkably high rates and the fit to these high rate offences is impressive.

While there are some discrepancies between the actual reoffending rates and the predicted rates (for example, child violence and indecent exposure, where the number of cases is very small), in general both models perform in an encouraging way. The goodness of fit statistics are all excellent, and show a good agreement between the observed and fitted rates.

6.3.3 Reoffending status by criminal history of target sanction

Table 11 presents actual and predicted reoffending percentages in terms of the criminal history of the target sanction. Actual reoffending rates for those with a single sanction are remarkably low, and appear to be lower for females. The fitted models overpredict reoffending rates for females and underpredict reoffending rates for males. The goodness of fit statistics for all four models are reasonably high and indicate a certain lack of fit for the one-year models.

Table 10. Reoffending status by type of target offence, comparing actual general reoffending rates with those predicted by the models.

		One-year reoffending			Two-year reoffending		
Target Offence	N	Actual %	Logistic %	Ordinal %	Actual %	Logistic %	Ordinal %
Violence	4361	0.280	0.274	0.277	0.407	0.406	0.401
Robbery	451	0.302	0.294	0.317	0.459	0.485	0.462
Public order-riot	3529	0.329	0.327	0.323	0.447	0.455	0.455
General sexual	115	0.183	0.263	0.243	0.313	0.355	0.356
Child sexual	228	0.118	0.091	0.100	0.158	0.174	0.166
Solicitation/ prostitution	133	0.496	0.479	0.455	0.556	0.549	0.569
Domestic burglary	645	0.434	0.432	0.462	0.673	0.659	0.613
Other burglary	646	0.525	0.494	0.504	0.689	0.665	0.642
Theft	5850	0.511	0.500	0.484	0.609	0.606	0.616
Handling	605	0.446	0.447	0.456	0.595	0.612	0.595
Fraud and forgery	923	0.275	0.290	0.266	0.372	0.371	0.385
Abscond from bail	689	0.601	0.603	0.609	0.746	0.754	0.740
Taking vehicle no consent	752	0.528	0.495	0.500	0.672	0.654	0.638
Theft from vehicle	289	0.571	0.515	0.510	0.678	0.657	0.646
Other motoring	2326	0.449	0.447	0.444	0.602	0.587	0.582
Drink driving	3317	0.169	0.163	0.157	0.248	0.252	0.252
Criminal damage etc	2427	0.333	0.329	0.328	0.473	0.467	0.463
Drugs import/ export/ prod	440	0.159	0.147	0.156	0.266	0.257	0.250
Drugs possess/ supply	3022	0.299	0.301	0.304	0.447	0.443	0.437
Other	931	0.310	0.283	0.278	0.408	0.388	0.389
Goodness of fit X2			15.98 on 19 df	21.42 on 19 df		5.23 on 19 df	13.14 on 19 df

Table 11. Reoffending status by length of criminal career, comparing actual general reoffending rates with those predicted by the models.

		One-year reoffending			Two-year reoffending		
Target sanction criminal history	N	Actual	Logistic	Ordinal	Actual	Logistic	Ordinal
1st caution/ warning/ reprimand	5302	0.157	0.153	0.147	0.246	0.252	0.254
2nd caution/ warning/ reprimand	1140	0.308	0.320	0.314	0.461	0.471	0.472
1st conviction	4239	0.110	0.099	0.088	0.169	0.158	0.160
other caution	1343	0.316	0.284	0.272	0.432	0.399	0.404
other conviction	19655	0.475	0.470	0.470	0.618	0.617	0.611
Goodness of fit X2			12.14 on 4 df	37.66 on 4 df		7.94 on 4 df	7.96 on 4 df

7. Model development

7.1 Exploring gender interactions

The analysis in the previous section, and in particular the examination of Table 9, suggested that there might be a separate age-crime curve for males and female offenders, and potentially different coefficients for other variables, such as the Copas rate.

We therefore proceeded to fit a series of additional models as follows:

a) the addition of an age by gender interaction
b) the fit of separate models to males and females.

a) The addition of an age by gender interaction.

We added the interaction term AGEBAND by GENDER to the analysis. The AIC value improved from 50611 to 50455 for the two-year reoffending model and from 49141 to 49045 for the one-year reoffending model. For the ordinal model, the AIC improved from 62432 to 62295, and the pseudo-AICs also improved to 49038 (one-year) and 50445 (two-year). The respective AUC values improved from 0.810 to 0.812 for two-year reoffending and from 0.796 to 0.797 for one-year reoffending. The AUC values for the ordinal model were 0.797 (one-year prediction) and 0.811(two-year prediction).

b) The fitting of separate models for males and females.

For the logistic models the AIC values further improved to 48984 (one-year), 50410 (two-year) over the age-gender interaction models above. AUC values further improved to 0.798 for the one-year model and stayed stable at 0.812 for the two-year model.

However, there were problems with some of the parameter estimates with this model because of small numbers of cases in some of the offending groups. In particular, there was a single female non-child sexual offender who was reconvicted between one and two years after the earlier conviction. This single case produced a large negative estimate for that category in the one-year model, and a large positive estimate in the two-year model. Similar problems occurred for the general sexual female offenders, none of whom were reconvicted before two years, giving a large negative estimate for the one-year model. While categories can be combined to remove this problem, this would ether lead to separate categories for males and females, or would mean a loss of offence group discrimination for the males. It would also involve the two sexual offending categories for females being merged with other offences, which is undesirable. Although we give the AUC values for this model in the table below, we do not consider it further.

7.2 Performance on the validation sample.

The table below gives the AUC values for these two additional models.

	One-year reoffending		Two-year reoffending	
	a) Ageband by gender interaction	b) Separate models for males and females	a) Ageband by gender interaction	b) Separate models for males and females
Logistic model	0.794	0.794	0.812	0.813
Ordinal model	0.794	0.794	0.812	0.812

The following table show the predicted one-year and two-year reoffending rates for the model including the age by gender interaction.

Current age		One-year reoffending			Two-year reoffending		
MALES	N	Actual %	Logistic %	Ordinal %	Actual %	Logistic %	Ordinal %
10 and under 12	164	0.250	0.309	0.299	0.402	0.449	0.451
12 and under 14	651	0.369	0.387	0.397	0.542	0.557	0.546
14 and under 16	1636	0.441	0.445	0.452	0.597	0.606	0.595
16 and under 18	2436	0.478	0.472	0.476	0.62	0.628	0.616
18 and under 21	4240	0.436	0.419	0.421	0.578	0.564	0.556
21 and under 25	4421	0.394	0.378	0.383	0.531	0.524	0.515
25 and under 30	3869	0.388	0.387	0.385	0.523	0.522	0.517
30 and under 35	3128	0.347	0.346	0.345	0.473	0.478	0.474
35 and under 40	2329	0.308	0.302	0.295	0.429	0.416	0.418
40 and under 50	2231	0.235	0.233	0.226	0.324	0.332	0.335
50 and over	1156	0.138	0.119	0.107	0.189	0.168	0.173
FEMALES							
10 and under 12	32	0.313	0.158	0.155	0.406	0.278	0.272
12 and under 14	239	0.218	0.240	0.222	0.326	0.340	0.353
14 and under 16	525	0.295	0.263	0.245	0.400	0.366	0.378
16 and under 18	501	0.257	0.289	0.267	0.345	0.379	0.392
18 and under 21	688	0.337	0.354	0.330	0.422	0.442	0.457
21 and under 25	804	0.367	0.364	0.348	0.461	0.457	0.470
25 and under 30	774	0.413	0.381	0.373	0.482	0.491	0.492
30 and under 35	668	0.323	0.327	0.308	0.416	0.416	0.427
35 and under 40	491	0.291	0.253	0.243	0.377	0.339	0.346
40 and under 50	489	0.221	0.207	0.194	0.282	0.289	0.289
50 and over	207	0.063	0.070	0.051	0.121	0.090	0.094
Goodness of fit X2			27.67 on 20 df	41.90 on 20 df		18.38 on 20 df	21.80 on 20 df

The table above shows a very good fit of expected reoffending to observed reoffending when cross-classified by age and gender, for both of the logistic models for one and two-year offending. It also shows an acceptable fit for the ordinal model in general. The one-year reoffending results for the ordinal model fit slightly less well, and this is caused by the poorer estimation for one-year reoffending for males aged over 50.

The equivalents of Table 10 and Table 11 for these additional models were both produced, with the results being virtually identical These extra tables have not been included in this report.

7.3 Discussion

The addition of an age by gender interaction has increased the AUC measures for both one-year and two-year reoffending, both for the logistic regression models and for the ordinal regression models. In addition, the fit to the age-gender table in the table above indicated a good degree of improvement over the results presented in Table 10. It provides a good compromise between the models presented in Section 5, and the fitting of completely separate models for gender, which present problems in estimation.

The improved fit comes at the cost of increased complexity in model fitting. If the logistic methodology is chosen, then there will in effect be four models – two for one-year offending (males and females) and two for two-year offending(males and females).

8. Conclusions

The aim of this report has been to develop a new measure to assess the performance of offenders following a sanction. This has involved a reappraisal of the current OGRS measure which has been widely used for assessing reconviction probabilities for groups of offenders. The OGRS measure has been remarkably successful in its basic task. However, the need for reappraisal – beyond the usual requirement to update a measure from time to time – comes about because of the new opportunities provided by the availability of the Police National Computer database over those given by the Offenders Index. The latter is a court-based database on convictions, whereas the former is an operational police system which records police activity as well as court information.

This opportunity to widen the analysis beyond conviction data encourages a reappraisal of the underlying philosophy of the measure. The new measure focuses on reoffending behaviour rather than a narrow focus on reconviction. It also widens the definition of offending by including a greater range of offences. While altering the analytical focus from reconviction to reoffending should appeal to criminologists, we need to stress that the outcome measure used is *proven* reoffending – proved either by conviction in court or by a caution, warning or other police sanction.

We maintain that this reconceptualisation is so fundamental that the new measure needs a new name, rather than simply being a variant of the OGRS measure. The shortcoming of the OGRS name for this new work is that it misleadingly places reconviction too prominently in the title. We suggest that the name ASPRO (<u>A</u> <u>S</u>core for <u>P</u>roven <u>R</u>eoffending of <u>O</u>ffenders) more clearly represents the nature of the work undertaken.

We have presented two possible score sets. The set based on logistic regression provides two separate scores – one for one-year reoffending and one for two-year reoffending. The second set based on ordinal Modelling, in contrast, provides a single score which can be used to predict both one and two-year reoffending.

Addition of an age by gender interaction term appears to improve the fit, and the extra work carried out in Chapter 7 appears beneficial. It is recommended that, despite the extra complexity, the age by gender interaction term be retained.

References

Copas, J. and Marshall, P. (1998). 'The Offender Group Reconviction Scale: the statistical reconviction score for use by probation officers.' *Journal of the Royal Statistical Society, Series C*, 47, 159-171.

Copas, J., Marshall, P. and Tarling, R. (1996) *Predicting reoffending for discretionary conditional release*. Home Office Research Study 150: Home Office

Criminal Records Bureau (2003) 'Weeding Rules'. *Disclosure News*, 9. Liverpool: Criminal Records Bureau, http://www.disclosure.gov.uk/newsletter/dn9/images/December%20DN.pdf

Francis, B. and Crosland, P. (2002) *The Police National Computer and the Offenders Index: can they be combined for research purposes?* Home Office: London Full report: http://www.homeoffice.gov.uk/rds/pdfs2/pncandoir170.pdf

Francis, B., Harman, J. and Humphreys, L. (2005) *Predicting Reconviction rates in Northern Ireland.* Research and Statistical Bulletin 7/2005. Northern Ireland Office. http://www.nio.gov.uk/index/statistics-research/publications.htm

Francis,B., Soothill, K., Humphreys, L. and Bezzina, A. (2005) *Developing Measures of Severity and Frequency of Reconviction.* Report submitted to the Home Office.

Harrell, F. Lee, K. and Mark, D. (1996) Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine*, 15, 361-387.

Hauck, W. and Donner, A. (1977) Wald's test as applied to hypotheses in logistic regression. *Journal of the American Statistical Association*, 72, 360, 851-3

Home Office (2004) *Criminal Statistics England and Wales 2003*. Cm 6361. Home Office: London.

Lesaffre, E. and Marx, B. (1993) Collinearity in logistic regression. *Communications in Statistics- Theory and Methods*. 22(7), 1933-1952

Lindsey, J. and Jones, B. (1998) 'Choosing among generalized linear models applied to medical data'. *Statistics in Medicine*, 17(1), 59-68.

Research Development and Statistics Directorate (1998). *The Offenders Index: Codebook*. Home Office: London. http://www.homeoffice.gov.uk/rds/pdfs/oicodes.pdf

Rex, S., Gelsthorpe, L., Roberts, C. and Jordan, P. (2004) *What's promising in community service: implementation of seven Pathfinder projects.* Home Office Research Findings 231: Home Office.

Stephens, K. and Brown, I. (2001) 'OGRS2 in Practice: An Elastic Ruler?' *Probation Journal*, 48(3), 179-187

Taylor, R (1999) *Predicting reconvictions for sexual and violent offences using the revised offender group reconviction scale*. Home Office Research Findings 104: Home Office.

Zhou, X, Obuchowski, N and McClish, D. (2002) Statistical Methods in Diagnostic Medicine. Wiley: New York.

Appendix A – Parameter estimates of various parametric models for reoffending.

APPENDIX A1 Estimates for logistic model for one-year and two-year reoffending. Final model, after backward elimination.

Parameter	One-year reoffending		Two-year reoffending	
	Parameter	Standard error	Parameter	Standard error
Targhist –1st caution/warn/rep	0		0	
Targhist –2ndt caution/warn/rep	0.1276	0.0637	0.0433	0.0584
Targhist –1st conviction	0.2384	0.0577	0.1452	0.0489
Targhist –other caution	0.4167	0.0679	0.2851	0.0625
Targhist –other conviction	0.4936	0.0480	0.4139	0.0441
Copas	1.2060	0.0219	1.3340	0.0229
Ageband 10 and under 12	0		0	
Ageband 12 and under 14	-0.0125	0.1468	0.0702	0.1363
Ageband 14 and under 16	-0.0124	0.1390	0.0418	0.1293
Ageband 16 and under 18	-0.1105	0.1383	-0.0631	0.1287
Ageband 18 and under 21	-0.5565	0.1375	-0.5425	0.1278
Ageband 21 and under 25	-0.9421	0.1382	-0.9244	0.1284
Ageband 25 and under 30	-1.0150	0.1387	-1.0330	0.1290
Ageband 30 and under 35	-1.1510	0.1396	-1.1830	0.1299
Ageband 35 and under 40	-1.2120	0.1414	-1.2770	0.1315
Ageband 40 and under 50	-1.3350	0.1423	-1.3920	0.1320
Ageband 50 and over	-1.8120	0.1557	-1.9970	0.1431
Gender –female	-0.1474	0.0325	-0.2728	0.0312
targoff violence	0		0	
targoff robbery	-0.6025	0.0958	-0.4790	0.0912
Targoff public order-riot	0.2163	0.0453	0.1476	0.0431
Targoff General sexual	0.1896	0.1916	0.0134	0.1813
Targoff Child sexual	-0.6957	0.2029	-0.5144	0.1621
Targoff solicitation/prostitution	0.9433	0.1727	0.6769	0.1761
Targoff domestic burglary	-0.1768	0.0759	0.0770	0.0802
Targoff other burglary	0.2577	0.0807	0.3219	0.0867
Targoff theft	0.7510	0.0404	0.5707	0.0397
Targoff handling	0.3517	0.0798	0.4217	0.0818
Targoff fraud and forgery	0.3276	0.0730	0.0799	0.0707
Targoff abscond from bail	0.7340	0.0749	0.7773	0.0829
Targoff taking vehicle no consent	0.4398	0.0748	0.4627	0.0783
Targoff theft from vehicle	0.5165	0.1182	0.4497	0.1267
Targoff other motoring	0.3096	0.0503	0.2349	0.0507
Targoff drink driving	-0.1030	0.0540	-0.1510	0.0490
Targoff criminal damage etc	0.2546	0.0499	0.2145	0.0477
Targoff drugs import/export/prod	-0.7672	0.1172	-0.7039	0.1011
Targoff drugs possess/supply	0.1009	0.0478	0.0928	0.0451
Targoff other	0.0082	0.0776	-0.0765	0.0743
cust4(1)	-0.1941	0.0512	-0.1946	0.0520
Constant	1.2030	0.1470	2.2390	0.1383

APPENDIX A2 Estimates for ordinal model for both one and two-year reoffending. Final model.

Parameter	One-year reoffending		Two-year reoffending	
	Parameter	Standard Error	Parameter	Standard Error
Threshold constants	1.1143	0.1360	1.8309	0.1362

Common parameters	Parameter	Standard Error
Targhist –1st caution/warn/rep	0.0000	0.0000
Targhist –2ndt caution/warn/rep	0.0967	0.0555
Targhist –1st conviction	0.1538	0.0476
Targhist –other caution	0.3590	0.0588
Targhist –other conviction	0.4766	0.0413
Copas	1.2529	0.0202
Ageband 10 and under 12	0.0000	0.0000
Ageband 12 and under 14	0.0463	0.1291
Ageband 14 and under 16	0.0266	0.1224
Ageband 16 and under 18	-0.0861	0.1217
Ageband 18 and under 21	-0.5533	0.1209
Ageband 21 and under 25	-0.9354	0.1215
Ageband 25 and under 30	-1.0300	0.1220
Ageband 30 and under 35	-1.1717	0.1228
Ageband 35 and under 40	-1.2449	0.1243
Ageband 40 and under 50	-1.3745	0.1249
Ageband 50 and over	-1.9475	0.1358
Gender –female	-0.2187	0.0289
targoff violence	0.0000	0.0000
targoff robbery	-0.5836	0.0843
Targoff public order-riot	0.1818	0.0401
Targoff General sexual	0.0229	0.1703
Targoff Child sexual	-0.6076	0.1590
Targoff solicitation/prostitution	0.8208	0.1580
Targoff domestic burglary	-0.1098	0.0701
Targoff other burglary	0.2520	0.0754
Targoff theft	0.6571	0.0363
Targoff handling	0.3680	0.0732
Targoff fraud and forgery	0.1778	0.0654
Targoff abscond from bail	0.7365	0.0709
Targoff taking vehicle no consent	0.4190	0.0692
Targoff theft from vehicle	0.4442	0.1105
Targoff other motoring	0.2579	0.0458
Targoff drink driving	-0.1395	0.0465
Targoff criminal damage etc	0.2270	0.0443
Targoff drugs import/export/prod	-0.7489	0.0971
Targoff drugs possess/supply	0.0939	0.0421
Targoff other	-0.0435	0.0687
cust4(1)	-0.1989	0.0467

§ variable omitted from final model.

APPENDIX A3 Estimates for logistic model for one-year and two-year reoffending. Alternative model with separate ageband parameters for males and females.

	One-year reoffending		Two-year reoffending	
	Parameter	Standard error	Parameter	Standard error
Targhist –1st caution/warn/rep	0		0	
Targhist –2nd^{t} caution/warn/rep	0.1109	0.0638	0.0331	0.0588
Targhist –1st conviction	0.2085	0.0577	0.1217	0.0490
Targhist –other caution	0.3956	0.0678	0.2774	0.0626
Targhist –other conviction	0.4729	0.0479	0.4057	0.0441
Copas	1.2114	0.0220	1.3390	0.0230
Male -Ageband 10 and under 12	0		0	
Male -Ageband 12 and under 14	-0.0036	0.1594	0.1337	0.1506
Male -Ageband 14 and under 16	0.0110	0.1498	0.1219	0.1419
Male -Ageband 16 and under 18	-0.1111	0.1482	-0.0046	0.1403
Male -Ageband 18 and under 21	-0.6456	0.1472	-0.5953	0.1388
Male -Ageband 21 and under 25	-1.0673	0.1478	-1.0226	0.1395
Male -Ageband 25 and under 30	-1.1427	0.1483	-1.1614	0.1401
Male -Ageband 30 and under 35	-1.3062	0.1494	-1.3209	0.1411
Male -Ageband 35 and under 40	-1.3177	0.1514	-1.3846	0.1429
Male -Ageband 40 and under 50	-1.4462	0.1524	-1.5053	0.1434
Male -Ageband 50 and over	-1.8907	0.1662	-2.0544	0.1545
Female - Ageband 10 and under 12	-0.8505	0.4207	-0.6906	0.3499
Female - Ageband 12 and under 14	-0.5716	0.1941	-0.6290	0.1793
Female - Ageband 14 and under 16	-0.6233	0.1661	-0.7000	0.1557
Female - Ageband 16 and under 18	-0.8681	0.1715	-1.0106	0.1614
Female - Ageband 18 and under 21	-0.8005	0.1628	-0.9652	0.1542
Female - Ageband 21 and under 25	-0.9710	0.1611	-1.0895	0.1531
Female - Ageband 25 and under 30	-1.0447	0.1627	-1.0459	0.1544
Female - Ageband 30 and under 35	-1.0480	0.1654	-1.1747	0.1569
Female - Ageband 35 and under 40	-1.3840	0.1742	-1.4475	0.1634
Female - Ageband 40 and under 50	-1.4704	0.1813	-1.5021	0.1682
Female - Ageband 50 and over	-2.1647	0.2767	-2.4616	0.2550
targoff violence	0		0	
targoff robbery	-0.6096	0.0962	-0.4860	0.0916
Targoff public order-riot	0.2125	0.0454	0.1432	0.0432
Targoff General sexual	0.2015	0.1916	0.0301	0.1815
Targoff Child sexual	-0.7179	0.2042	-0.5317	0.1632
Targoff solicitation/prostitution	0.8750	0.1760	0.6218	0.1801
Targoff domestic burglary	-0.1862	0.0760	0.0640	0.0803
Targoff other burglary	0.2452	0.0807	0.3023	0.0868
Targoff theft	0.7490	0.0405	0.5728	0.0398
Targoff handling	0.3345	0.0799	0.4040	0.0821
Targoff fraud and forgery	0.3066	0.0731	0.0562	0.0707
Targoff abscond from bail	0.7294	0.0750	0.7744	0.0831
Targoff taking vehicle no consent	0.4020	0.0750	0.4147	0.0788
Targoff theft from vehicle	0.4980	0.1181	0.4205	0.1267
Targoff other motoring	0.3111	0.0504	0.2363	0.0508
Targoff drink driving	-0.0943	0.0541	-0.1380	0.0491
Targoff criminal damage etc	0.2337	0.0501	0.1876	0.0479
Targoff drugs import/export/prod	-0.7794	0.1171	-0.7167	0.1010
Targoff drugs possess/supply	0.0876	0.0479	0.0756	0.0452
Targoff other	-0.0047	0.0775	-0.0938	0.0742
cust4(1)	-0.1771	0.0513	-0.1695	0.0521
Constant	1.3249	0.1559	2.3283	0.1483

APPENDIX A4 Estimates for ordinal model for one-year and two-year reoffending. Alternative model with separate ageband parameters for males and females.

	One-year reoffending		Two-year reoffending	
	Parameter	Standard Error	Parameter	Standard Error
Threshold constants	1.4199	0.1386	2.1393	0.1388

	Parameter	Standard error
Targhist –1st caution/warn/rep	0	0
Targhist –2ndt caution/warn/rep	0.0799	0.0557
Targhist –1st conviction	0.1258	0.0476
Targhist –other caution	0.3407	0.0588
Targhist –other conviction	0.4605	0.0413
Copas score	1.2581	0.0202
Male -Ageband 10 and under 12	0	0
Male -Ageband 12 and under 14	0.0827	0.1416
Male -Ageband 14 and under 16	0.0746	0.1333
Male -Ageband 16 and under 18	-0.0617	0.1318
Male -Ageband 18 and under 21	-0.6251	0.1307
Male -Ageband 21 and under 25	-1.0485	0.1313
Male -Ageband 25 and under 30	-1.1592	0.1318
Male -Ageband 30 and under 35	-1.3166	0.1327
Male -Ageband 35 and under 40	-1.3527	0.1344
Male -Ageband 40 and under 50	-1.4837	0.1350
Male -Ageband 50 and over	-2.0071	0.1460
Female - Ageband 10 and under 12	-0.7842	0.3453
Female - Ageband 12 and under 14	-0.6140	0.1703
Female - Ageband 14 and under 16	-0.6706	0.1469
Female - Ageband 16 and under 18	-0.9617	0.1520
Female - Ageband 18 and under 21	-0.8994	0.1448
Female - Ageband 21 and under 25	-1.0315	0.1434
Female - Ageband 25 and under 30	-1.0543	0.1446
Female - Ageband 30 and under 35	-1.1283	0.1471
Female - Ageband 35 and under 40	-1.4186	0.1537
Female - Ageband 40 and under 50	-1.5243	0.1591
Female - Ageband 50 and over	-2.4469	0.2442
targoff violence	0.0000	0.0000
targoff robbery	-0.5908	0.0846
Targoff public order-riot	0.1772	0.0402
Targoff General sexual	0.0368	0.1703
Targoff Child sexual	-0.6323	0.1599
Targoff solicitation/prostitution	0.7548	0.1610
Targoff domestic burglary	-0.1188	0.0702
Targoff other burglary	0.2382	0.0754
Targoff theft	0.6575	0.0364
Targoff handling	0.3503	0.0733
Targoff fraud and forgery	0.1552	0.0654
Targoff abscond from bail	0.7334	0.0710
Targoff taking vehicle no consent	0.3779	0.0695
Targoff theft from vehicle	0.4224	0.1105
Targoff other motoring	0.2599	0.0459
Targoff drink driving	-0.1290	0.0466
Targoff criminal damage etc	0.2023	0.0444
Targoff drugs import/export/prod	-0.7608	0.0970
Targoff drugs possess/supply	0.0788	0.0421
Targoff other	-0.0599	0.0687
cust4(1)	-0.1791	0.0468

APPENDIX B1 Creating a baseline against which to compare the new ASPRO model

We wanted to create a baseline against which the predictive power of ASPRO can be compared.

We did this by fitting a pseudo-OGRS2 model to the PNC data used for the reoffending study. In other words variables were created akin to those that are used in the OGRS2 model but using PNC equivalents. For example whereas OGRS2 uses 'age at first *conviction*' the pseudo-model will use 'age at first *sanctioning occasion*'. Whereas the OGRS2 model uses '*reconviction* within 2 years' as the response variable the pseudo-model will use '*reoffending* within 2 years'.

OGRS2 (Taylor, 1999) uses the following variables

1. offender's age in years at time of sentence/current conviction (this was split into ten age bands)
2. gender
3. number of youth custodial sentences
4. current offence group (This was based on only standard list offences and broke down offences into 27 detailed offence categories)
5. age at first conviction
6. the Copas rate variable
7. history of burglary (whether the offender has a current or previous history of burglary)
8. History of breach (whether the offender has a current or previous history of a breach).

Table B1(a) gives details of the variables used for this task

Table B1(a) List of variables used in Pseudo-OGRS2 model

	Variable name	**Notes**	**Coding**	**Explanatory/response**
1.	AGEBAND2	Using the same 10 bands as OGRS2	Under 14 = 1, 14 and under 16 = 2, 16 and under 18 = 3, 18 and under 21 = 4, 21 and under 25 = 5, 25 and under 30 = 6, 30 and under 35 = 7, 35 and under 40 = 8, 40 and under 50 = 9, 50 and over = 10	Explanatory
2.	GENDER		Male = 1, Female = 2.	Explanatory
3.	PREVCUST2	This is number of **sanctioning occasions** in which a custodial sentence was received in which the offender is under 21. Not surprisingly, the values of this variable		Explanatory

		are the same as if only **court appearances** were used.		
4.	TARGOFF2	The same 27 offence groups	1 = Violence, 2 = Robbery, 3 = Aggravated burglary, 4 = Violence against children, 5 = Public order riot offences, 6 = Firearm offences, 7 = Sexual offences, 8 = Sex offences against children, 9 = Indecent exposure, 10 = Solicitation /prostitution/homosexual offences, 11 = Domestic burglary, 12 = Non-domestic burglary, 13 = Other burglary, 14 = Theft, 15 = Handling, 16 = Fraud and forgery, 17 = Absconding and bail offences, 18 = Taking and driving away and related offences, 19 = Theft from cars, 20 = Other motoring, 21 = Criminal/malicious damage, 22 = Drugs (import/export/production), 23 = Drugs (supply), 24 = Drugs (possession allowing on premises), 25 = Drugs (possession & supply), 26 = Arson, 27 = Other	Explanatory
5.	FIRSTAGE2	This was translated into **age at first sanction**	Under 14 = 1, 14 and under 16 = 2, 16 and under 18 = 3, 18 and under 21 = 4, 21 and under 25 = 5, 25 and under 30 = 6, 30 and under 35 = 7, 35 and under 40 = 8, 40 and under 50 = 9, 50 and over = 10	Explanatory
6.	COPAS2	For OGRS2 this is calculated by calculating the square root of the number of court appearances divided by length (time from age at first conviction to age at target) +5. The pseudo-model calculated this by calculating the square root of the **number of sanctioning occasions** divided by **length (time from age at first sanction to age at target**) +5.		Explanatory

7.	ANYBURG	For OGRS2 this is a binary variable indicating at least one conviction for burglary. For the pseudo-model this binary variable indicates **at least one sanction for burglary** (the same variable used for the new model)	No = 0, yes = 1	Explanatory
8.	ANYBRC	For OGRS2 this binary variable indicates **at least one conviction for breach**. For the pseudo-model this binary model indicates **at least one sanction for breach** (the same variable used for the new model)	No = 0, yes = 1	Explanatory
	RECON1	Proven reoffending within 1 year (the same variable used for the new model)	No = 0, yes = 1	Response
	RECON2	Proven reoffending within 2 years (the same variable used for the new model)	No = 0, yes = 1	Response

The models for one and two-year reoffending were fitted to the same calibration dataset using logistic regression. As was done with the main analysis, these models were validated using the 40% of the sample not used to fit the model.

Table B1(b) Area under the ROC curve (AUC) for the competing frameworks

	One-year reoffending	Two-year reoffending
Logistic model Pseudo-OGRS2	0.790	0.809
Logistic model New model (Chapter 5)	0.793	0.811
Logistic model New model (Chapter 7)	0.794	0.812

As table B1(b) shows there is an improvement in fit in moving from the OGRS2 framework to the new models.

We also considered calculating ‘true’ OGRS2 parameters to predict re-offending on our PNC dataset. The parameters produced by Ricky Taylor for OGRS2 were calculated using OI data and are used to predict reconviction and so this is not an appropriate route to take. For example, the outcome variable used in the OI based OGRS2 is “Standard List reconviction” -

whereas the outcome variable we have been using from PNC data is proved reoffending. In addition, the predictor variables in OGRS2 are based on Standard List convictions - the PNC holds information on all forms of sanctioning and so the predictor variables created from PNC data would be calculated on previous sanctions of all kinds. Nevertheless it would be possible to create pseudo OI variables. So would recalibrating the OGRS2 model on pseudo-OI data achieve anything? Well, it would show how well a model used to predict reconviction from reconviction data performs in the task of predicting reoffending using reoffending data - but only if the equivalent/pseudo (OI) predictor and outcome variables were calculated from our data.

Appendix C Detailed definition of TARGOFF in terms of Home Office Codes

Offence Category	Home Office Codes
Violence.	1.00 to 4.03, 4.05, 4.07 to 9.99, 11.00 to 15.99, 29 to 29.99, 31.00 to 31.99, 35 to 36.99, 79.02 to 79.99, 81.00 to 81.99, 103.00 to 105.99, 109.00 to 109.99, 111.00 to 111.99, 113.00 to 113.99, 15.00 to 15.99, 145.00 to 145.99
Robbery	34.00 to 34.99
Public Order/Riot	64.00 to 66.99, 125.00 to 125.48, 125.57 to 125.58, 132.00 to 132.99, 140.00 to 141.01, 141.03 to 141.07, 141.10 to 141.99, 162.00 to 162.99, 164.00 to 164.07, 182.00 to 183.99, 502.00 to 502.99
Sexual	6.00 to 16.11, 16.13 to 16.16, 16.18 to 16.22, 16.29 to 16.99, 17.00 to 17.04, 17.06 to 17.10, 17.12, 17.13, 17.15, 19.00 to 19.06, 19.08, 19.10, 19.12, 19.14, 20, 20.02 to 20.03, 20.05, 22.02 to 22.05, 23.02, 23.06 to 23.07, 23.10 to 23.13, 23.16 to 23.17, 23.20 to 23.21, 23.30 to 23.37, 25.00 to 25.02, 25.04 to 25.99, 70.00 to 70.99, 73.00 to 73.06, 86 to 86.01, 86.03 to 86.09, 88.03 to 88.11, 139 to 139.99, 164.12, 503.00 to 503.99, 504.00
Sexual child	16.12, 16.17, 16.23 to 16.28, 17.05, 17.11, 17.14, 17.16, 19.07, 19.09, 19.11, 19.13, 19.16 to 19.19, 20.01, 20.04, 20.06, 21.00 to 21.99, 22.00, 22.06 to 22.25, 23.01, 23.03 to 23.05, 23.08 to 23.09, 23.14 to 23.15, 23.18 to 23.19, 23.22 to 23.29, 25.03, 71.00 to 71.99, 73.07 to 73.18, 74.00 to 74.02, 86.02, 86.10, 88.01 to 88.02, 192.01, 192.02
Soliciting/prostitution	18.00 to 18.00, 24.00 to 24.99, 27.00 to 27.99, 107.00 to 107.99, 165.00 to 167.99, 187.00 to 187.99
Domestic Burglary	28.00 to 28.99
Other Burglary	30.00 to 30.99, 32 to 33.99
Theft	38.00 to 38.99, 39 to 44.99, 46.00 to 47.99, 49.00 to 49.99, 118.00 to 118.99, 119.00 to 119.00, 174.00 to 174.99
Handling/Receiving Stolen goods	54.00 to 54.99, 120.00 to 120.99, 178.00 to 178.99
Fraud to Forgery	50.00 to 50.99, 51 to 53.99, 55.00 to 55.99, 60.00 to 61.99, 154.00 to 154.99
Absconding/Bail Offences	80.00 to 80.99, 83.00 to 83.99
Taking Mot, Vehicles to Driving Away to Related offences	37.00 to 37.99, 48.00 to 48.99, 130.00 to 130.02, 131.00 to 131.99
Theft From Mot, Vehicle	45.00 to 45.99
Other Mooting Offences	4.04, 126.00 to 129.99, 130.03 to 130.04, 130.06, 170.00 to 170.99, 173.00 to 173.99, 802.00 to 802.99, 804.00 to 805.99, 807.00 to 807.99, 809.00 to 970.03
Drink/Drugs Driving	4.06, 130.05, 141.02, 141.08 to 141.09, 803.00 to 803.99

Criminal/Malicious Damage	56.00 to 59.99, 149.00 to 149.99, 185.00 to 185.99
Drugs Import/Export/Production etc	77.01, 77.02, 77.19, 77.20, 77.50, 92.01 to 92.49
Drugs Possession/Supply etc	77.00, 77.03, 77.04, 77.06, 77.09 to 77.49, 77.51 to 77.60, 92.50 to 92.89, 93.00 to 93.99, 193 to 193.99
Other	All other codes not included above