

## DAL\_2018\_0001\_IOM\_Model

### Questions raised at the last Ethics Committee:

- 1. Can the underlying variables and patterns being identified be described more clearly, so there is a greater understanding of what kinds of offences may be a reliable predictor of later more serious offences, and what variables are present in explaining that trajectory and why?**

The majority of the features used for the model relate to criminal history (number of crimes over a period, number of times in custody and the RFSDi harm prior to transition to the high harm group as well as raw harm as calculated using the Cambridge crime harm index (CCHI) and ONS crime severity score (over certain periods of time), etc. Other features relate to how central individuals are in networks and if mentioned in intelligence logs, what those logs referred to. For a data dictionary of the features used, etc., please see appendix A to this note.

Appendix A also shows some of the patterns in the resulting model by way of showing (for the top 5 most important features) how moving through their values (as in moving from 0 – 100 for example) present in the data affect the resultant estimated probabilities.

It should be noted that the model is currently undergoing a rebuild following changes in data availability and processes as well as a change to the stop and search feature. This may lead to some minor differences in feature importance, etc. in the final model.

- 2. Can more details be provided about who the target cohort is including offence types and how harm is defined?**

For the RFSDi, the cohort for analysis is everyone on the crimes system who has a designation of defendant – i.e. there has been enough evidence for WMP to charge the individuals in that particular instance. This does not include any individuals who have only been, for example, suspects or victims, etc. For individuals who have been both charged and suspects (in separate crimes), only those crimes with which they have been charged are included.

For the predictive modelling element, again, the predictions are made for each individual within the crimes system where they have been charged (and only offences for which they have been charged are taken into account).

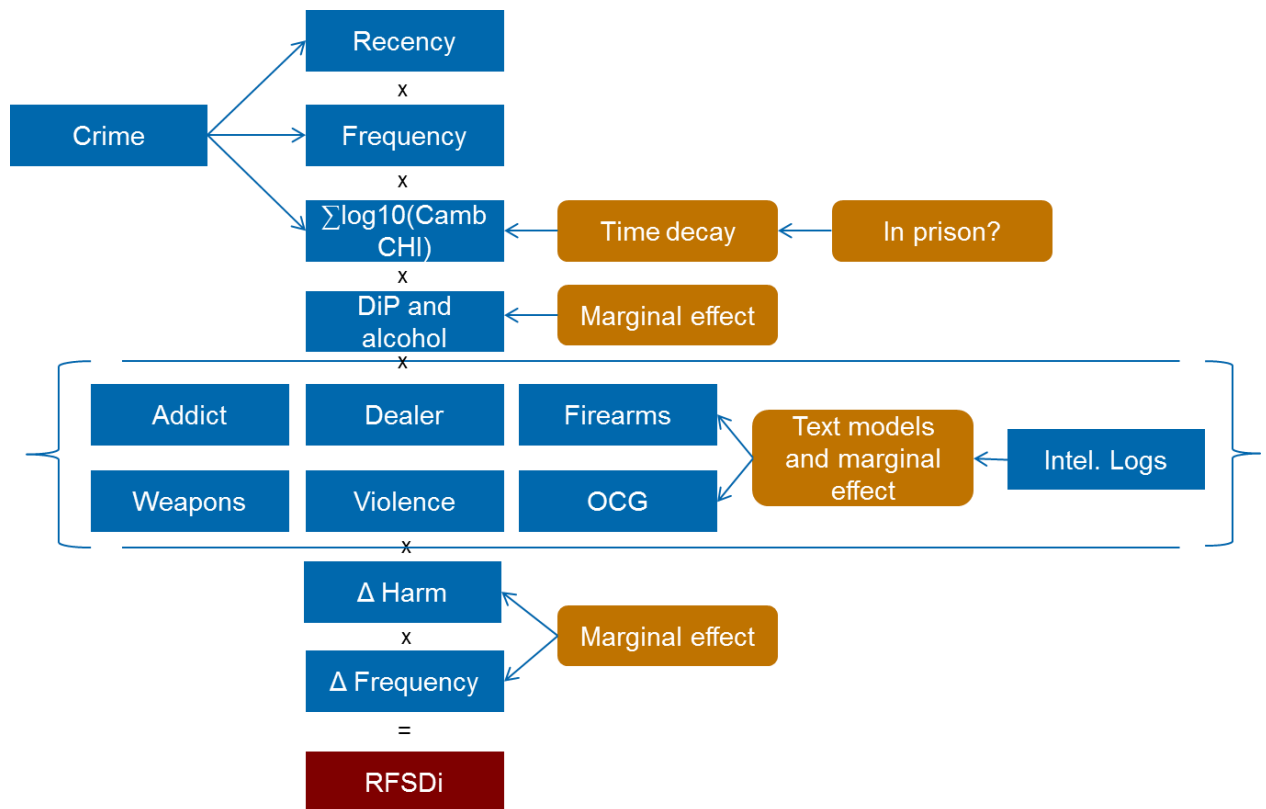
The ‘target’ cohort in this instance would be either those individuals who have a high RFSDi score and aren’t currently managed (subject to the offender managers’ assessments), or those individuals who aren’t currently in the high harm groups, but they are more likely than not to become high harm in the near future with a view to Offender Managers examining each case to ascertain appropriateness of management; the predictive element essentially providing a means of an initial filtering, a direction of who could be initially assessed.

Any decision to place an individual within the offender management process would be decided upon by offender managers.

Harm is defined via calculating the RFSDi with, for the purposes of the model, low harm designated as 0 (the low, low middle and high middle groups) and high harm, the 1s, being the high and super high groups.

RFSDi is Recency, Frequency, Severity, Drugs and intelligence.

The broad process for the construction of the RFSDi is outlined in the diagram below:



The crime history of individuals is ordered and individuals are allocated to 1 of 10 relative bins in terms of the recency of their crime and the frequency (essentially the average number of crimes per year). These measures are relative over the individuals, so recency is in terms of the recency over the whole dataset rather than some hard break point.

Each of the crimes is then matched to a figure from the Cambridge Crime Harm Index (CCHI). This is based on the guideline number of days' sentence for a first offence for particular crimes. The log to the base 10 is then actually used in order to reduce the potential for volatility and skewing of the score due to the orders of magnitude differences between the different crimes. This is then summed over each individual following the application of an exponential weight decay to reflect the passage of time (so, for example, a robbery committed 2 years ago scores lower than a robbery committed yesterday). SMEs considered that it would be important to reflect the harm of those who are currently serving custodial sentences without this being reduced due to time effects. For this reason, for those who are currently in prison, the time decay starts from the point at which they entered prison (for the current sentence), essentially the RFSDi is calculated for these individuals as if their sentence start date were the current date. For any crimes committed post the start of their

sentence, the RFSDi is calculated as for everyone else (the start date for the weight decay is the current day).

Other measures of crime harm / severity such as the Crime severity score published (experimentally) by the Office for National Statistics were also investigated. The ONS score, when applied to individuals, correlates highly with the CCHI (spearman correlation of circa 0.8 and a maximum information criterion of 0.8). It was decided, at this stage, to use the CCHI due to its democratic accountability.

When consulting with SMEs it was also considered that the use of drugs and alcohol were important considerations in the committing of crime and so in creating harm. Whether a crime was related to alcohol or whether a nominal is addicted to alcohol has been taken from the crimes system. Drug usage (and whether a nominal has tested positive for a particular substance) has been taken from data arising from the drug intervention programme (DIP) data. These have then formed weights to be applied as part of the RFSDi. The weights have been determined via statistical techniques (by regressing the centred  $\log_{10}(\text{CCHI})$  on the types of crimes committed, years active in crime and the feature of interest). The resultant weights therefore are essentially some fraction of the standard deviation of the  $\log_{10}(\text{CCHI})$  which are then applied.

SMEs also considered it important to include information from the intelligence logs as to drug dealing, addiction, firearms, violence, weapons (other than firearms) and affiliation with organised crime groups (OCGs). Whether a nominal is mentioned in an intelligence log that refers to these subjects has been ascertained via statistical techniques (naïve Bayes models (with document type frequency as a prior) using document term matrices as the features which identifies whether a log refers to these issues). Logs have been linked to individuals via the GID creation process and so weights are applied (and calculated in the same fashion as for alcohol addiction, etc. with a further weight (reducing the initial weight's value) then being applied.).

Changes in behaviour are also accounted for in the RFSDi by way of applying weights (ascertained via a similar method as for drug addiction, firearms, etc.) in relation to the trajectory of an individual's behaviour in terms of harm and frequency – if the trajectory is upwards, a positive weight is applied, if downwards, a negative weight (the weights being derived in a similar fashion to the use of firearms, etc.).

The different elements are then multiplied to produce the final RFSDi.

**3. Can more detail be provided about data reliability including an assessment of why data is not considered too old, or how, for example, the assessment of intelligence data is deemed reliable and does not inadvertently label someone by association with other people's criminality?**

Changes in some datasets are apparent over time in terms of some patterns (e.g. the occurrence of some crimes) which reflect changes in law, changes in reporting practices and changes in policy / priorities as well as changes in the general criminal environment.

Although data within the crimes system go back some 20 years, only data from the last 10 years have been used. This is because the majority of the features used in building the model are based on crimes (where they have been charged) covering a 2 year period over the last 8 years; this period being required to generate a reasonable number of observations where an individual has crossed from low or middling levels of harm to high levels of harm (this being the target variable of interest).

In relation to the features used, the crimes data do not show any great changes over this time period.

In relation to the intelligence data, two main steps are taken to ensure the veracity of the data, firstly only intelligence that meets the following requirements as set out in the national intelligence grading scheme; it is (a) reliable and (b) it is either known directly or if indirectly has been corroborated. Furthermore, the weights applied resulting from intelligence as part of the RFSDi are themselves down weighted so as to reduce their importance in the overall RFSDi calculation.

It is also the case that as a result of the management of police information (MoPI) guidelines, some data need to be retained for a very long period of time. The length of retention essentially comes down to the history of the individual (minor crimes that would otherwise be deleted need to be retained if an individual has committed more serious offences which of themselves would be retained for a long period of time). The retention (or not) of data also applies to intelligence data. The finalised version of the model will be built using MoPI(ed) data.

Generally intelligence that is older tends to be less relevant in most cases than newer information, for example telephone numbers and addresses. However certain types of information, such as the suspicion of sexual offences, are likely to be relevant for a person's life (due to MoPI). In terms of association between two people, it depends on the nature of that relationship rather than because there is one at all. For example having a brother does not mean that both brothers are criminals but it may be relevant in the short term. But if the information suggests that they commit crime together then is it relevant for longer.

Whilst the grading scheme is utilised for the purposes of the model, in its own right the intelligence grading scheme does not help determine which information is more relevant for longer. It suggests what needs to be done with the information before it is suitable to be acted upon. For example an unreliable intelligence log may suggest an immediate threat to life which needs to be acted upon as soon as possible, but a reliable intelligence log about a stolen car does not need immediate action. It is the content of the information that determines how long it is relevant for. Information about sexual offences and other violence will be relevant for a very long time, while other criminality such as burglary may not be relevant for long at all.

**4. How is the model going to be used operationally and what will be the benefit to policing purposes?**

The model seeks to address a capacity issue and allows the assessment of more data (than previously utilised) related to individuals in order to identify those people who require enhanced support and offender management. This cannot be done manually and as a result opportunities are currently missed to help offenders by providing preventative interventions and also protect future victims from harm.

It is perhaps the way the results of the model are used that is important. Starting from the most risky and harmful offenders, an intelligence or offender management professional needs to assess each result from the model and decide the best course of action (which may be no action) based on an intelligence led approach. Each individual identified by the model requires a bespoke approach and be provided with support that suits their individual needs. The model will not be used to generate a list of offenders that are then handed out for “action” as this would be counterproductive. For example an offender in the County Lines method of criminality may also be a vulnerable adult who has been coerced into crime. While they are an offender, they are also a victim who has needs that need to be met to prevent them from causing further harm to themselves and others.

The benefits of using the model in the right way is that there will be fewer occurrences of harmful or vulnerable offenders being missed. They are more likely to receive the support they need, and the professional assessment of the model will ensure that the planning and coordination of the support with police and partners is appropriate.

**5. Far more detail is required around what interventions might be applied to those individuals identified, bearing in mind that potential adverse consequences of inaccurate predictions will be largely dependent on the type of intervention carried out, and as regards associated policies and procedures to ensure all relevant information taken into account and weighted appropriately.**

In June 2009 the Government published the Integrated Offender Management Policy Statement. Its purpose was to provide all criminal justice partnerships with direction and support in bringing together the management of repeat offenders into a more coherent structure.

Integrated Offender Management is the strategic umbrella and overarching framework that brings agencies together to prioritise interventions with offenders causing harm to their communities. IOM built upon and compliments previously separate offender programmes such as PPO, MAPPA and DIP.

Under the current system, Offender Management Teams (OMT) will manage those identified as HIGH RISK (via the existing Corvus score) with the support from specialist force resources and Neighbourhood Teams (NT) manage those identified as MEDIUM RISK with support of the OMT.

Currently, it is acknowledged that the assessing of risk posed by offenders will require both a review of data and consultation with partners is crucial in order to come to a shared

understanding of the risk. There will generally be variation between partner agencies in terms of their assessment of the risk posed by particular offenders, but there will be consistent identification at Force level and delivering a local response to local problems.

While partners have different assessments of risk, these should be factors in decisions made regarding offenders and include information sharing, co-located multi agency hubs and ODOC and these should inform and support decisions to take subjects to an appropriate tasking meeting.

Integrated Offender Management (IOM) looks to case manage offenders away from offending and towards better lifestyle choices using a range of tactics based on a model of “control” and change”. The IOM teams have referral pathways in place with a range of statutory, voluntary, and private providers based on the 7 pathways of reduced reoffending:

- Accommodation;
- Education, employment and training;
- Mental and physical health;
- Attitude, thinking and behaviour;
- Children and families;
- Drugs & Alcohol; and,
- Money and benefits issues.

There are also 2 further pathways bespoke to women offenders, Domestic Abuse and Sexual Health.

**6. It was suggested that future meetings would benefit from additional contextual explanations regarding the use of risk assessment and prediction in offender management and safeguarding.**

Much of the risk assessment undertaken by Offender Managers is based around the current Corvus score as noted above. Other information that comes into play however includes factors such as their accommodation status, employment status, drug usage, alcohol usage, associates, whether there is an imminence of re-offending, etc.

At present therefore, whilst there is not a formal means of prediction, in managing the risk posed by individuals, Offender Managers are implicitly making predictions with those that enter into their purview as to their likelihood of re-offending, if so in what ways, their likelihood of creating increased levels of harm in the future and (taking individuals receptiveness to interventions into account), the likelihood of success.

**7. How is effectiveness going to be measured, both in terms of model accuracy and outcomes achieved from supportive interventions?**

For beta testing (whereby the model(s) are productionised and the results of the first round of predictions are compared to new data over a period of 2 – 3 months), the same measures of accuracy (sensitivity, specificity, precision, area under the curve, etc.) will be assessed. These measures will be compared to those obtained during model build (from application to the test dataset) but also compared to random; the predictions should be at least 20% better than random.

As part of the on-going productionised process, predictions will be made against an older version of the dataset so that measures of predictive accuracy can be ascertained (predominantly the ROC curve) so that we can identify if the model's predictive accuracy is degrading.

**8. Will the use of the model by WMP officers / staff be obligatory?**

It is not currently envisaged that use of the model by OMs will be obligatory.

**9. How will the tool be incorporated into the overall decision making process? Far more detail is required on the balance between who will use the model, when and how it fits in with their professional discretion.**

Whilst the RFSDi score would (eventually) be made available to many within WMP, the outputs from the predictive model element will be used and assessed by offender management professionals. In the case of the RFSDi score, for example, intelligence professionals could include the score in their assessments to provide insight into current or potential future offending, as they currently do, when recommending which individuals and actions to prioritise to prevent offending by using the intelligence led daily, weekly and monthly tactical tasking processes. In most cases this will be to help address the current force level and local priorities such as robbery or knife crime. Working with the offender management team, intelligence professionals will also recommend activities to prevent serious crime including sexual offences even if they sit outside force and local priorities.

It is envisaged that unexpected results will be assessed in the same way as all results and an intelligence led decision will be made to decide what activity needs to be done next. It could be no activity, further data collection or immediate action to protect the offender or a potential victim from harm.

**10. As ethnicity has been removed as a factor from the model, what other proxy measures might there be for ethnicity, such that the model risks perpetuating ethnic stereotypes or confirming historic biases?**

In order to check for this we have run a (multinomial) regression of the features against ethnicity (i.e. a statistical model where ethnicity is the target variable and the features used are the same as the features used in the model). Essentially this is an examination as to whether anything correlates with ethnicity.

All of the estimated coefficients are near zero which indicates that none of the features identify ethnicities. See appendix B for details (for ease of reference, rather than the raw coefficients, the charts show the odds).

**11. Who will have access to the information, where will the outputs be communicated and stored, and is there a potential for this to cause bias?**

Whilst the RFSDi may be reported to response via the Insight Search application, it is intended that Offender Managers and their supervisors would have access to the RFSDi and the outputs of the predictive model(s) via a dashboard. This allows for self-serve reporting (e.g. counts by NPU, etc.) and for viewing the outputs at the levels of the individual (which Offender Managers need to know in order to make an assessment as to whether and in what way individuals should be managed).

It is not envisioned that any information arising from the RFSDi or the predictive model(s) will be communicated outside of the Offender Management department (with the exception of the RFSDi to response).

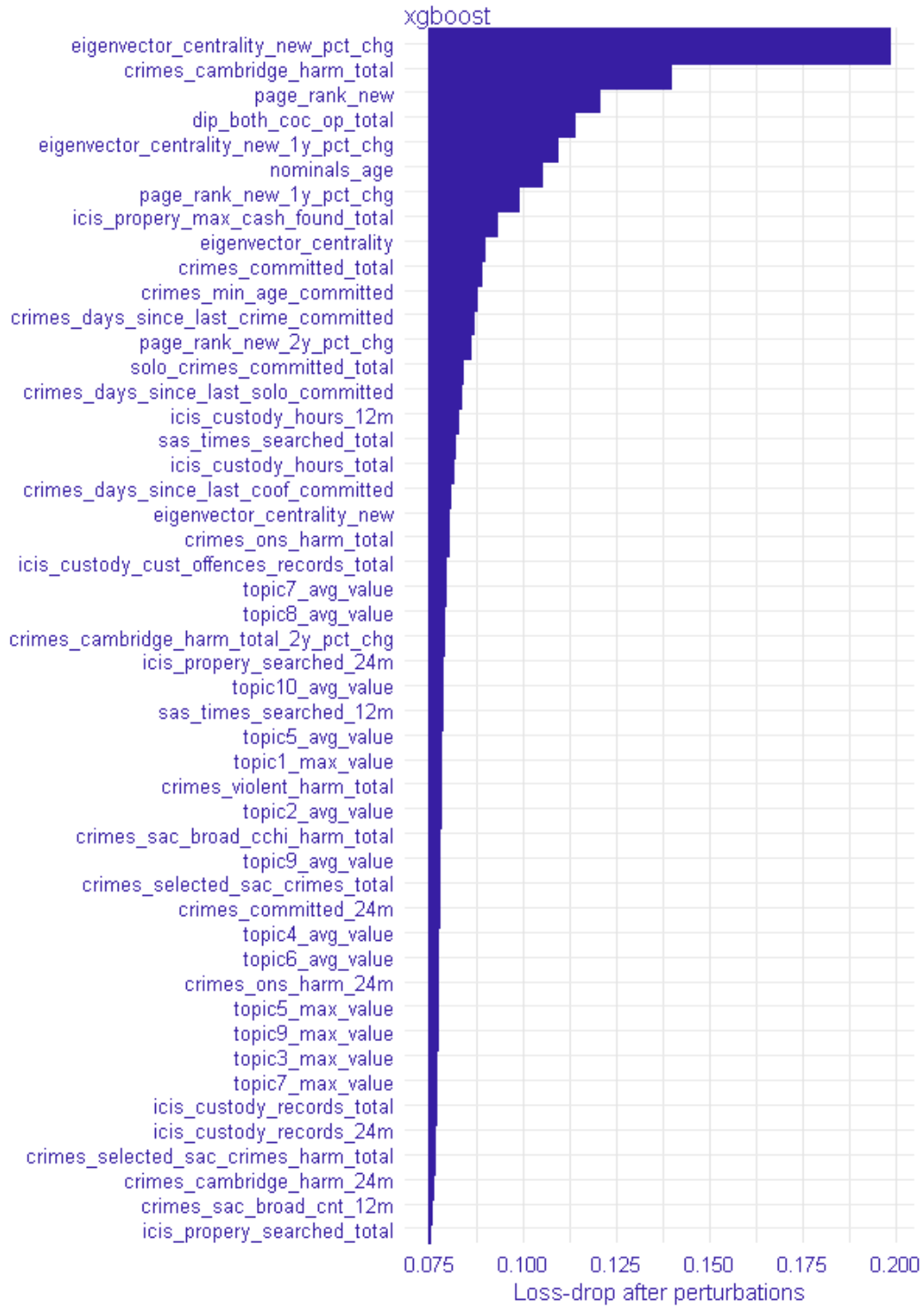
Any process of risk assessment, statistical or not, predictive or not will have the potential to lead to biases of some form. It is intended that the use of the dashboard will be tracked, this will enable us to identify the ways in which the dashboard is used and by whom and so, coupled with the outcomes of the Offender Management process for individuals, assess the pathway from prediction through to outcome which will enable an assessment of the presence of any biases in usage. Of course this will take time for the system to be used (and data to be received).

We have met with the individual Local Offender Management Units (LOMUs) and introduced them to the underlying aims, approaches and a dashboard mock-up (to receive their feedback). Prior to productionisation (as a later part of the beta testing phase), it is also intended to use a small group of Offender Managers to test the dashboard. This will be followed by training for the individual LOMUs prior to go live (including the requirement for non-communication of results outside of the department).



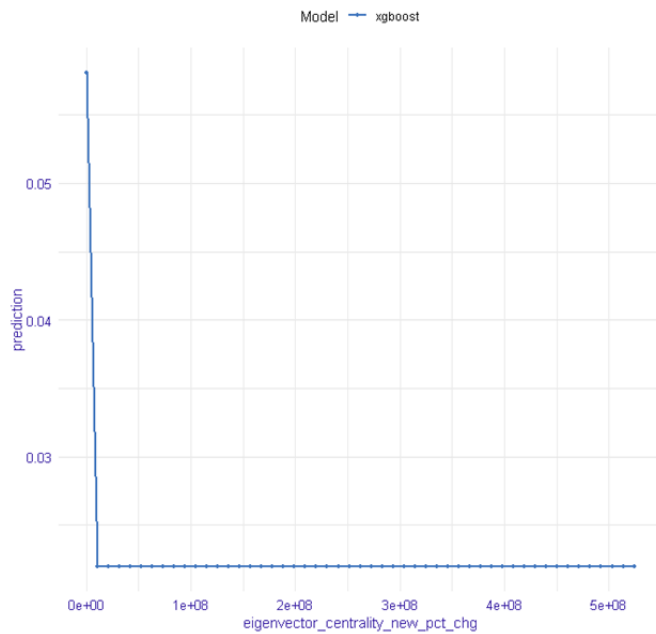
APPENDIX A – QUESTION 1:

The chart below shows the feature importance (in order):

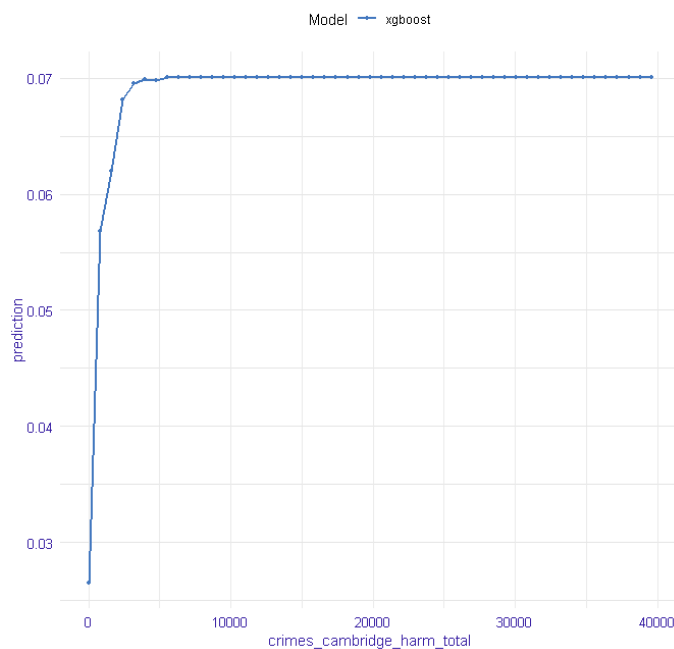


## Patterns:

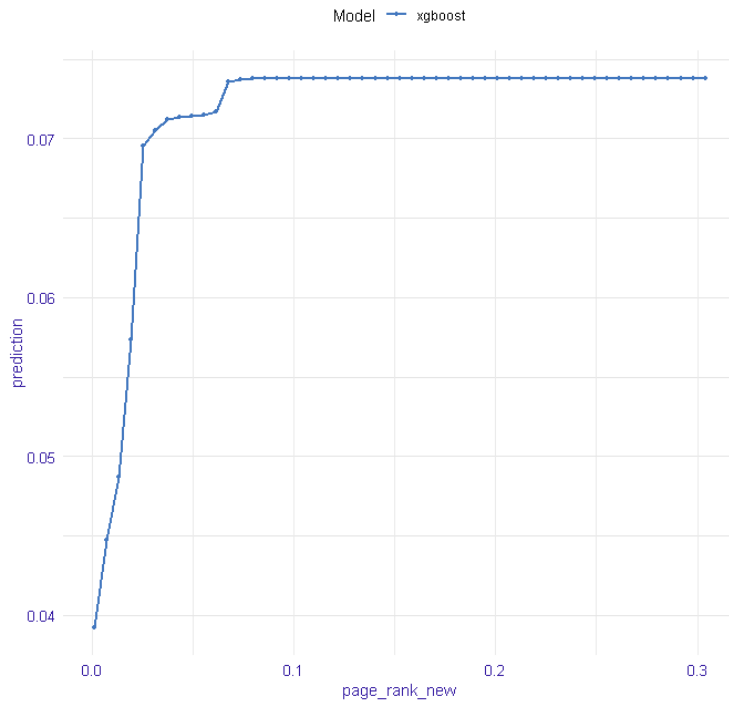
The charts below show the general effect of the top 5 features on the estimated probabilities resulting from the model.



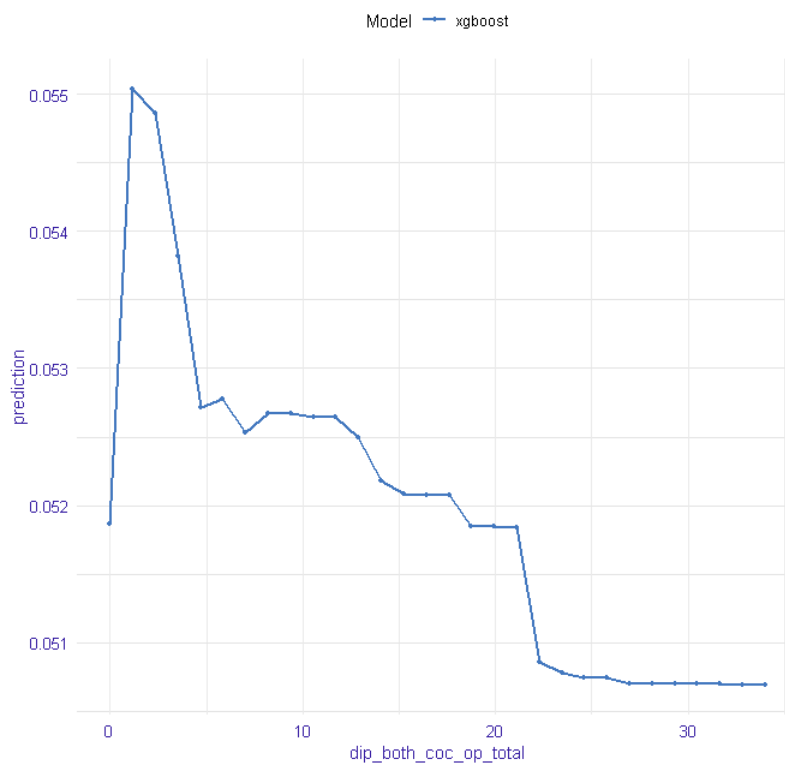
The change in eigenvector centrality essentially has a negative relationship with the probability of becoming high harm. This may reflect a pattern of small (absolute) changes (as the measure is percentage change) either leading to promotion through a network leading to less crime being committed or it may reflect that early promotion leads to large changes in the measure, but they are not as yet committing more serious crimes.



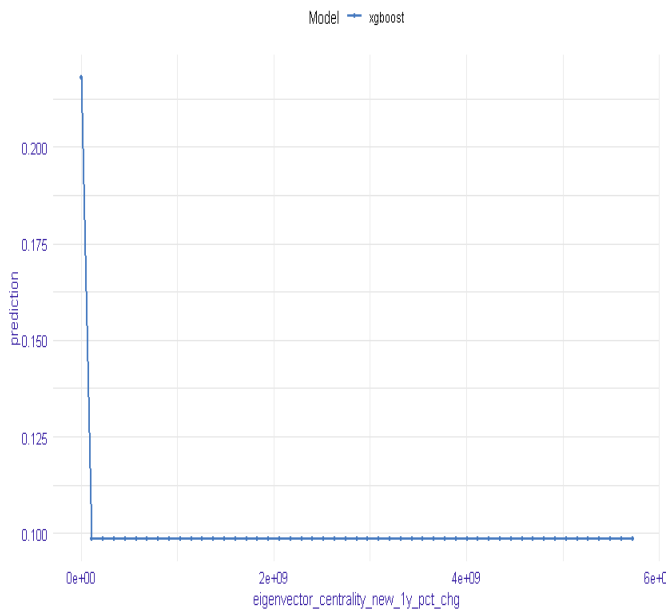
This ('raw') measure of harm moves positively with the estimated probability in that the more harm they have committed, the more likely they are to transition.



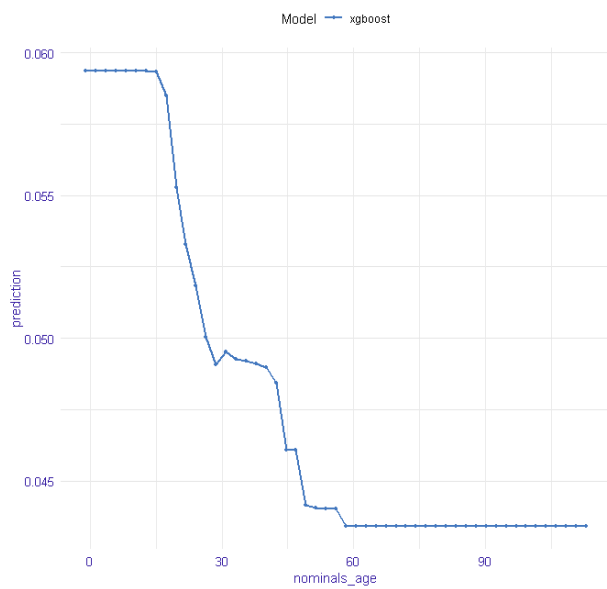
This measure of (current) centrality has a positive relationship with the estimated probability in that the more central someone is in a network, the more likely they are to become high harm.



This is a more ambiguous pattern compared to the other features; initially, the more positive tests, the more likely they are to become high harm. However, if this continues the less likely they are to become high harm. This may reflect a pattern that although they are still taking drugs, being within a structured programme means they are less likely to continue creating harm through criminal activity.



This is a similar pattern to the 2 year percentage change above.



The age of individuals moves negatively with the probability of becoming high harm – the older the individual, the less likely they are to become high harm.

Feature Dictionary:

crimes_cambridge_harm_24m	Total CCHI harm caused over last 24 months
crimes_cambridge_harm_total	The total harm as measured by the CCHI so far
crimes_cambridge_harm_total_2y_pct_chg	% change in CCHI harm over 2 years
crimes_committed_24m	Number of crimes committed over 24 months
crimes_committed_total	Total number of crimes committed so far
crimes_days_since_last_coof_committed	Number of days since last crime committed with a co-offendent
crimes_days_since_last_crime_committed	Days since last crime committed
crimes_days_since_last_solo_committed	Number of days since last solo (by themselves) crime committed
crimes_days_since_last_solo_committed2	Number of days since last solo (by themselves) crime committed (a slightly different measure)
crimes_min_age_committed	Minim age at which a crimes was recorded for an individual
crimes_ons_harm_24m	Total ONS severity score over last 24 months
crimes_ons_harm_total	The total harm as measured by the ONS crime severity score so far
crimes_sac_broad_cchi_harm_total	Total CCHI harm caused via SAC crimes (broad definition)
crimes_sac_broad_cnt_12m	Number of broad SAC offences over last 12 months
crimes_selected_sac_crimes_harm_total	Total harm created via serious aquisitive crimes (SAC)
crimes_selected_sac_crimes_total	Number of various SC offences
crimes_violent_harm_total	Total harm created through violent crimes
dip_both_coc_op_total	Number of tests found positive for cocaine via DIP
eigenvector_centrality	A measure of the centrality of in individual in a network (1 year ago)
eigenvector_centrality_new	A measure of the current centrality of in individual in a network
eigenvector_centrality_new_1y_pct_chg	A measure of how central a person is in a network (% change over 1 year)
eigenvector_centrality_new_pct_chg	A measure of how central a person is in a network (% change over 2 years)
icis_custody_cust_offences_records_total	Total number of ICIS records
icis_custody_hours_12m	Number of hours in custody over last 12 months
icis_custody_hours_total	Total number of hours in custody so far
icis_custody_records_24m	Number of custody records over last 24 months
icis_custody_records_total	Total number of custody records
icis_propery_max_cash_found_total	Total amount of cash found in property searches
icis_propery_searched_24m	Number of times a property search is recorded in ICIS system over 24 months
icis_propery_searched_total	Number of times a property search is recorded in ICIS system
nominals_age	Age
page_rank_new	A measure of the current centrality of in individual in a network
page_rank_new_1y_pct_chg	A measure of how central a person is in a network (% change over 1 year)
page_rank_new_2y_pct_chg	A measure of how central a person is in a network (% change over 2 years)
sas_times_searched_12m	Number of times stopped and searched in last 12 months
sas_times_searched_total	Total number of stop and searches
solo_crimes_committed_total	Total number of solo crimes committed
topic1_max_value	Maximum probability of pertinent IMS logs being topic 1

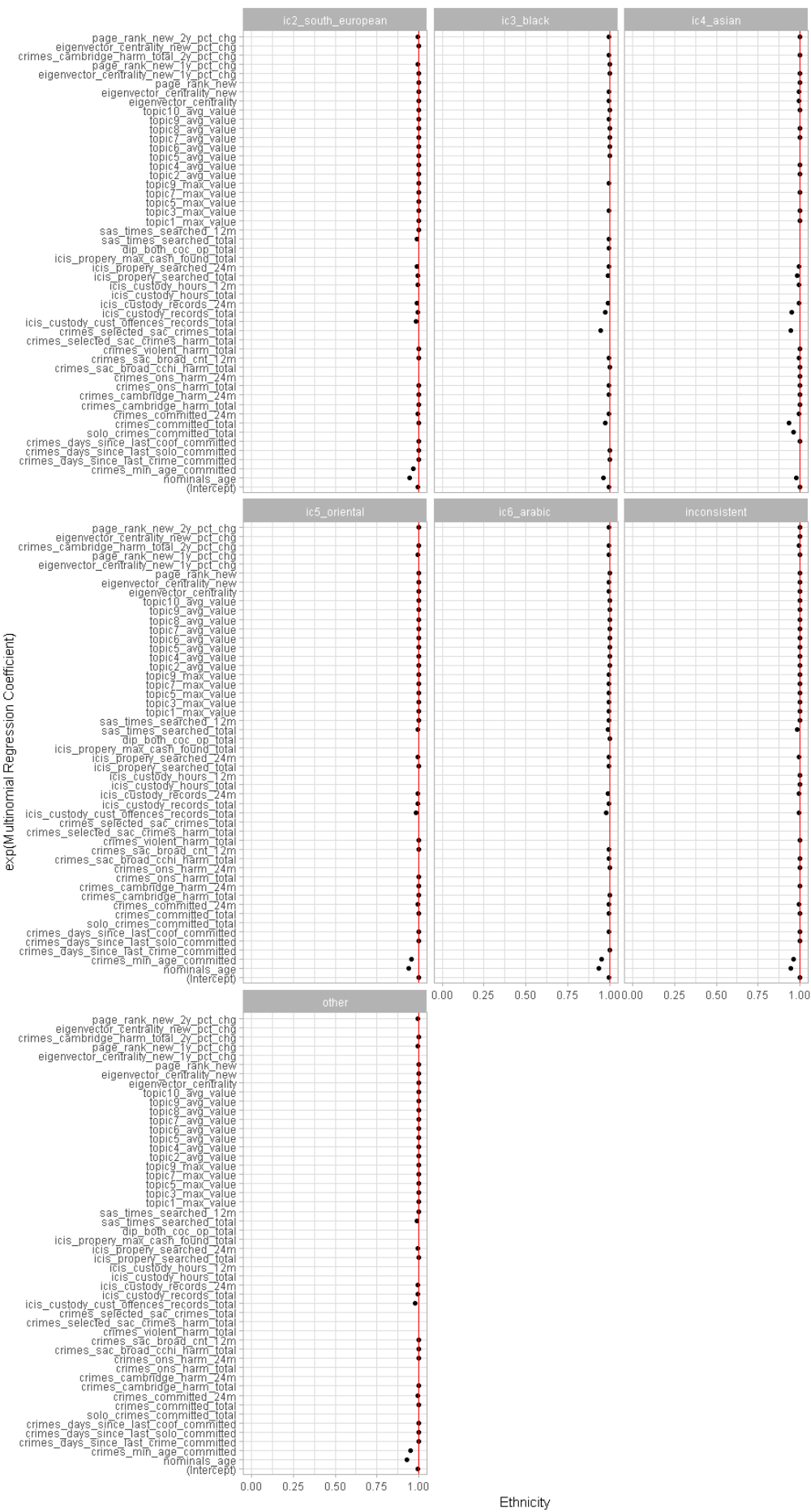
topic10_avg_value	Average probability of pertinent IMS logs being topic 10
topic2_avg_value	Average probability of pertinent IMS logs being topic 2
topic3_max_value	Maximum probability of pertinent IMS logs being topic 3
topic4_avg_value	Average probability of pertinent IMS logs being topic 4
topic5_avg_value	Average probability of pertinent IMS logs being topic 5
topic5_max_value	Maximum probability of pertinent IMS logs being topic 5
topic6_avg_value	Average probability of pertinent IMS logs being topic 6
topic7_avg_value	Average probability of pertinent IMS logs being topic 7
topic7_max_value	Maximum probability of pertinent IMS logs being topic 7
topic8_avg_value	Average probability of pertinent IMS logs being topic 8
topic9_avg_value	Average probability of pertinent IMS logs being topic 9
topic9_max_value	Maximum probability of pertinent IMS logs being topic 9

#### APPENDIX B – QUESTION 10:

The charts below show the odds of being the ethnicity in question compared to being IC1 – North European given whatever the feature in question is. If the probability was the same as for IC1 North European, the odds would be 1 (a probability of 50% or 50/50).

For example; given the `crimes_min_age_committed`, the odds of being IC2 – Southern European rather than IC1 – North European are 0.95 which means that the probability of being IC2 rather than IC1 given this feature is 0.49.

It can be seen from all the charts that almost all the odds are the same as for IC1 – North European.



Ethnicity