2015

Analytical Services Branch, DOE NIRSS to 2020 - Developing a novice driver indicator

The purpose of the paper is to document the methodology behind the novice driver indicator reported in the NIRSS to 2020.

Methodology Paper

Analytical Services Branch
DOE

Updated September 2015

Contents

Introduction

Each year Analytical Services Branch (ASB) in DOE publish their annual statistical report which presents the most recent position against targets and key performance indicators (KPIs) as set out in the NIRSS to 2020[1]. The report measures change from the previous year and the strategy baseline period (2004-2008)[2]. There were four key performance indicators which were not monitored in previous reports – one of which related to novice drivers. There was no agreed method to report information for this indicator so in the absence of an agreed benchmark ASB were tasked to explore options and develop a method to enable reporting. The indicator which this work relates to is:

- Novice driver casualties: deaths and serious injuries within 6, 12 and 24 months of passing test.

Investigation Stage

The data sources initially identified to feed into the novice driver indicator were the Police Service of Northern Ireland (PSNI) road traffic collisions data and Driver and Vehicle Agency (DVA) licensing data. The PSNI dataset for drivers involved in collisions where there was a fatal or serious casualty holds, in the majority of cases, a driver license number; it is important to be aware that this particular field is not validated so there are some accuracy and completeness issues with regard to this data item.

ASB initially carried out a data matching exercise to link PSNI and DVA licensing files, this work highlighted a number of more significant issues with regard to the licensing data primarily to do with the accuracy of the 'date passed test' field. Having talked through these issues with DVA, it was decided that DVA test data, which also contains licence number, would be a more reliable source of information on when a candidate passed their driving test and hence should be matched with the PSNI collision dataset.

Methodology

Beginning with the PSNI collision data for the required analysis period (i.e., 2008-2014), ASB identified the vehicles (cars, taxis and LGVs) that were involved in KSI collisions. From this file all the vehicles which appeared to have a valid driver licence number, and hence were capable of being matched, were then selected (approx 75% of all vehicles involved in road traffic collision). Whilst this is not ideal, it should not affect the resultant novice driver estimates as long as the achieved c75% sample is still representative of all novice drivers,

[1] http://applications.doeni.gov.uk/publications/document.asp?docid=19681
[2] http://www.doeni.gov.uk/index/information/asb/statistics/road_safety_statistics.htm#roadsafetystrategy

i.e., the cases with missing/invalid licence numbers are purely random and not alike on some key characteristic such as length of driving experience.

This sample was then matched, using the licence number field, with the test data from DVA (containing all category B test passes from 2006 to 2014) enabling date of birth, gender and 'category B' test date from the DVA to be appended onto the selected PSNI records. Note that it was not practical to obtain, and then attempt to match, on the full DVA test file. The selected time period was chosen as anyone with a test pass date earlier than 2006 would have to have been driving for a minimum of 2 years before the KSI analysis period which began in 2008. A key assumption was made, therefore, that any collision record that could not be matched in the restricted DVA test dataset was because the driver had at least 2 years driving experience, post-test, at the time of the collision. This assumption was tested using a random sample of non-matched collision records and was shown to be robust.

On completion of various validations and quality checks with DVA, ASB were satisfied that the match was the optimum that could be achieved. ASB then carried out checks to ensure that the profile of the group with matched driver licence numbers (in terms of age and gender) was not obviously different to the population. These checks did not highlight any systematic bias in the achieved sample.

The time period between the date the collision occurred and the driver pass test date was calculated making it possible to report for this indicator as set out in the NIRSS to 2020. All the work that has been described above was based solely on vehicle data so it was necessary to build in the casualty data to report on the number of deaths and serious injuries as detailed in the indicator description. Due to the outputs being based on a sample of cases, it was also necessary to weight the results back to NI levels. Confidence intervals were produced alongside the final estimates so that users could gauge their robustness.

Concluding Remarks

This paper documents at a high level the methodology that ASB have applied to develop a data source that can monitor and report against the 'Novice driver casualties: deaths and serious injuries within 6, 12 and 24 months of passing test' KPI.

ASB decided to report KSI casualties relating to all collisions in which a novice driver was involved with a split between those collisions for which the novice driver was deemed responsible and those for which they were not. Note that it was not possible to produce

results by single year due to an insufficient number of cases, therefore, three years data has been combined to report annual averages.

ASB have opted to move from the within 6, 12 and 24 month of passing test and reporting the indicator by the time periods stated below.

- 0-6 months
- 7-12 months
- 13-18 months
- 19-24 months
- 0-24 months.