Enhancing Data in National Crash Statistics using an In-depth Study.

Jonas Krampe, Mirko Junge

Abstract After successfully introducing projects to reduce traffic accident deaths, the European Union focuses on lowering the seriously injured as defined by the Abbreviated Injury Scale (AIS). We used in-depth crash investigation data for injury severity projection to get MAIS3+ (Maximum AIS) injury severity distribution information for a national data set.

In contrast to the projection method by weighting, the approach presented here maps the distribution of interest, i.e., the distribution of a particular variable, from the in-depth study to the national level. The scheme works by transferring conditional probabilities—estimated using the in-depth survey—to the national level and using these conditional probabilities to supplement the distribution information from the national database. Even though the enhanced national data set provides only injury severity distribution information, it can be used to analyse the temporal stability of the crash environment at different injury severity levels.

As an application of the method, we added MAIS2+, 3+, and 4+ injury severity distribution information, using the German in-depth Accident Study (GIDAS) information, to the German national crash data set and investigated the shift in injury severity over the crash years. Each new crash test standard shows a statistically significant injury severity reduction.

Keywords Projection, data enhancement, national crash statistics, in-depth crash investigation, MAIS3+ injury severity level.

I. INTRODUCTION

Policy evaluation, e.g., evaluating new advanced driver assistance systems or monitoring political objectives such as *Vision Zero* of the European Union [1], is a significant application of crash databases. However, national crash databases often lack detailed information for such assessments; consequently, biased in-depth crash studies are often used for policy evaluation. In-depth crash studies such as National Automotive Sampling System Crashworthiness Data System (NASS-CDS) for the US or German in-depth Accident Study (GIDAS) for Germany are helpful for policy evaluation only if projected to the national level, which means the information of the in-depth study needs to be transformed such that it is representative on the national level.

One approach to projecting an in-depth crash study to a national level is weighting the in-depth cases. The in-depth crash study is treated as a sub-sample of the national crash data. Therefore, each case in the national crash data has a probability of being included in the in-depth crash study. Under the assumption of non-zero inclusion probabilities, the weights can be chosen as their inverse. For instance, the weights in NASS-CDS are determined this way. To elaborate, NASS-CDS uses a multi-stage sampling scheme for case selection. In the first step, representative counties and police jurisdictions are chosen. Then, the researchers select a sub-sample for an in-depth investigation of the crashes occurring in the selected region. Finally, the selection scheme is chosen so that more severe injuries are more likely to be sampled, and the focus is on tow-away crashes only [2].

In NASS-CDS the selection probabilities at each stage are known. Thus, the product of all selection probabilities results in the inclusion probabilities for each case; inversion of the inclusion probabilities determines the weights [3]. Note that the inclusion probability of non-tow-away crashes is zero. Thus, the weighted NASS-CDS sample can only represent all US tow-away crashes and not all US crashes. In the present study, we denote such an approach as weighting by inclusion probability.

The in-depth study GIDAS, see Section III (In-depth Sample (GIDAS)) for further details, is a sub-sample of German crash data. However, GIDAS does not follow a statistical sampling scheme, and the probabilities for case inclusion are unknown. Instead, GIDAS has a specific sampling criterion: A crash with an injured participant [4].

J. Krampe is a postdoc research fellow in the Statistics and Data Science Department at Cornell University, New York and iTUBS, Braunschweig, Germany (jonas.krampe@cornell.edu). M. Junge is Senior Researcher in the Department of Accident Research at Volkswagen Group Innovation, Wolfsburg, Germany.

References [5-6] proposed to project these samples to the national level by weighting the in-depth study by the weights chosen such that the contingency tables for a set of covariates matches. We denote this approach as weighting by matching of covariates. For instance, possible covariates are crash location or crash type. Since too many covariates lead to an overfit, the covariates must be carefully selected. Reference [6] suggest evaluating the fit using a crash severity measure available in both data sets, such as the police-reported crash injury severity. However, since the weights are built upon covariates available in both databases, i.e., national level and in-depth, it is indeterminable how well this approach works for variables only available in the in-depth study. Due to the sample criterion, non-severe crashes are underrepresented in GIDAS, and most likely, some crash types are not included. Thus, some inclusion probabilities are zero, making the computation of the universal weights impossible—universal in the sense that a weighted in-depth study transfers representatively all variables of the in-depth study to the national level. In particular, universal weights require that the in-depth data set be split into different subgroups so that each in-depth subset is representative of the corresponding subgroup on the national level.

A sample criterion such as the one used in GIDAS requires very particular subgroups. For example, for crashes such as 40-year-old men in full-frontal single-vehicle crashes with $\Delta v=40$ km/h, the in-depth study and the national database would be similar even though the in-depth study tends to more severe crashes. However, if an important variable such as age or Δv is dropped, a sample criterion leading to more severe crashes would mean that the in-depth study is not representative anymore. Due to limited crash information in the national data set, such detailed subgroups cannot be built using the national database. Hence, the weights cannot be calculated. Additionally, the sample sizes of the different subgroups would be tiny, which could make statistical reasoning difficult. Thus, it is indeterminable if such universal weights can be achieved by matching contingency tables of covariates.

This paper goes beyond the weighting with universal weights by changing the perspective: Instead of weighting the in-depth study to match the distribution of some covariates at the national level, it enhances the information in the national database with data from the in-depth study. In contrast to the previous projection methods by weighting, the approach presented here maps the distribution of interest, i.e., the distribution of a particular variable, from the in-depth study to the national level. We denote this as mapping of conditional probabilities. That means the projection is not made case-by-case but on an aggregated level, which can be understood as a (sub)-population level. The approach works by transferring conditional probabilities—estimated using the in-depth study—to the national level and using these conditional probabilities to enhance the national database. The underlying assumption is that the used conditional probabilities coincide between both data sets. Thus, if the assumption holds, the enhanced national database possesses the same representative properties as the original national database. In this case, the enhanced national database can be considered an unconditional data set, e.g., not restricted to tow-away crashes (NASS) or crashes with at least one injured occupant (GIDAS). Such unconditional national data is beneficial if the interest is on trend in the injury severity distribution over time, as sampling criteria like *crashes with at least one injured person* could lead to a non-negligible bias. Furthermore, working on an aggregated level, not on a case-by-case level, has the advantage that only aggregated information is required to do the fit. For instance, no case-level information from the national database is needed, which can be an advantage regarding privacy requirements. The method is illustrated in Figure 1. A more detailed mathematical description is given in Section II (METHODS).

Let us elaborate on the main argument of transferring conditional probabilities. This approach assumes that the conditional probabilities coincide between the in-depth study and the national database so that information can be transferred. This assumption may not hold for both data sets in general but for the particular sub-populations of interest. From the perspective of the weighting approach, conditional probabilities can be translated to weights. However, if the national database is enhanced with multiple variables, different weights would be used for each variable. Hence, the method presented here can be seen as a generalisation of the weighting approach where different weighting schemes might be used. Additionally, working with conditional probabilities has the advantage of interpreting conditional probabilities is easier than weights alone. Both mapping of conditional probabilities and weighting by matching of covariates rely on non-testable assumptions. However, the interpretability of conditional probabilities can help in arguing whether it is likely or not that the conditional probabilities of the in-depth study coincide with the national database for the particular sub-population of interest.

Note that this approach is not limited to using a single in-depth study, i.e., the national database can be enhanced with multiple variables where the information for each of these variables might come from different in-depth studies.

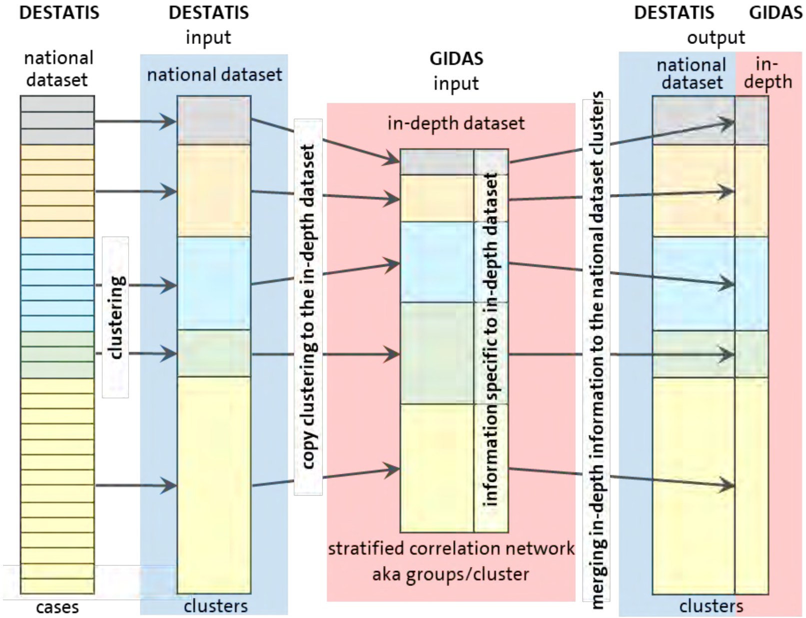

Fig. 1: Supplementing the marginal distributions of the national data set (blue) with the in-depth data set (red).

Returning to our outset, the mean objective of the European *Vision Zero* is to reduce the number of fatalities and serious injuries as defined by the Maximum Abbreviated Injury Scale (MAIS3+) [7] to zero on European roads by 2050 [1]. However, many national crash databases do not provide AIS injury information, making close monitoring difficult. The mapping of conditional probabilities approach can help to fill this gap. This paper presents an application of mapping of conditional probabilities to German crash data. Hence, it closes the monitoring gap for *Vision Zero* in Germany and shows how new insights can be gained from the enhanced national data set. The DESTATIS (Deutsches Statistik-Informationssystem, engl. German statistic information system) national crash data is supplemented with detailed medical data from an in-depth GIDAS database. This approach enables the investigation of the shift in injury severity over crash years. One striking finding of such investigation is that the crash environment exhibits notable temporal instability even after controlling for factors like model vehicle year and focusing exclusively on injury severity of occupants in modern vehicles. Specifically, it is observed that the nature of traffic accidents is changing from year to year, becoming increasingly aggressive. This temporal instability underscores the dynamic and evolving nature of road safety challenges, see among others [8-9].

II. METHODS

This section defines the method *mapping of conditional probabilities* and elaborates on its mathematical details.

Data Enhancement

Let X and Y be discrete random variables with the outcomes $x_0, \dots, x_p$ and $y_0, \dots, y_q$, respectively. Furthermore, the in-depth study has information on X and Y, whereas the national database contains information on X only. We mark by subscript N the national level and by subscript I the in-depth study, i.e., Y_I marks a variable in the in-depth study. Note that the distribution of Y_N is unknown, and the objective is to enhance the national level with information about Y_N. The approach is as follows. First note that we have for any outcome $y_i \in \{y_0, \dots, y_q\}$

$$\mathbb{P}(Y_N = y_i) = \sum_{j=0}^{p} \mathbb{P}\left(Y_N = y_i | X_N = x_j\right)\mathbb{P}\left(X_N = x_j\right). \tag{1}$$

On national level the probabilities $\mathbb{P}\left(X_N = x_j\right)$ are known (or can be estimated at least). Then, to enhance the national level with $\mathbb{P}(Y_N = y_i)$ the underlying assumption is that

$$\mathbb{P}(Y_N = y_i | X_N = x_j) = \mathbb{P}(Y_I = y_i | X_I = x_j) \quad \text{for all } i, j. \tag{2}$$

Hence, we obtain

$$\mathbb{P}(Y_N = y_i) = \sum_{j=0}^{p} \mathbb{P}\left(Y_I = y_i | X_I = x_j\right) \mathbb{P}(X_N = x_j). \tag{3}$$

That is, the previously unknown distribution of Y_N is now given by known (or at least estimable) quantities. To apply this to data, the probabilities $\mathbb{P}(X_N = x_j), j = 0, \dots, p$ can be estimated using the national data, and the conditional probabilities $\mathbb{P}(Y_I = y_i | X_I = x_j), i = 0, \dots, q, j = 0, \dots, p$ can be estimated using the in-depth data.

Note the assumption, i.e., Equation 2, may not generally hold for the conditional probability of Y given X. The idea is first to build subgroups that can be considered homogeneous in both data sets, so the assumption holds. For instance, some crucial covariates like occupant age or traffic domain may affect the distribution of Y and the conditional distribution of Y given X. Hence, the approach is first to stratify the data for these covariates such that the assumption holds for all subgroups. To elaborate, let Z be a covariate with outcomes $z_0, \dots, z_r$, and the covariate Z is observed at the national level and in the in-depth study. Then, assumption Equation 2 can be modified to

$$\mathbb{P}(Y_N = y_i | X_N = x_j, Z_N = z_k) = \mathbb{P}(Y_I = y_i | X_I = x_j, Z_I = z_k) \quad \text{for all } i, j, k \tag{4}$$

and the distribution of Y_N is then given by

$$\mathbb{P}(Y_N = y_i) = \sum_{k=0}^{r} \sum_{j=0}^{q} \mathbb{P}\left(Y_I = y_i | X_I = x_j, Z = z_k\right) \mathbb{P}(X_N = x_j, Z = z_k). \tag{5}$$

Stratification of conditional probabilities is only necessary in dimensions in which the distribution of the stratification variable differs among the data sets.[1]

Since the data enhancement uses the distributional level only, only data stratified with respect to the variables X and Z (in the in-depth study also for Y) is necessary, but nothing further. Thus, the involved probabilities can be estimated using aggregated data, and no case-by-case data is required, which can be a potential benefit in terms of privacy constraints.

III. APPLICATION

As an application to the methodology, an example is discussed in detail, emphasising the importance of covariates to build homogeneous subgroups is illustrated.

At the beginning of the new millennium, the European Union (EU) instigated projects to reduce road traffic fatalities by 50 % within a decade [10]. However, even though the programme was renewed in 2010 for another 50 % reduction in 10 years, an influential publication by [11] shifted the focus from road traffic death to the severely injured, i.e., MAIS3+ injured [7].

The national statistics do not code injury severity using the *AIS*-scale. Still, the injury severity documented by police officers is based on treatment location, duration, and survival (henceforth denoted as P-scale, see Section III (The P-Scale)) and does not provide enough detail to answer questions about severely injured persons, i.e., the MAIS3+ injured. The approach *mapping of conditional probabilities* is used to enhance the national crash data for drivers of passenger vehicles with information on injury severity, focusing on severely injured persons. Note that the evaluation is for drivers of motor vehicles only, as the national crash data does not provide seating position information for other vehicle occupants.

As an application of the method, timelines of crashes are investigated: The influence of the vehicle model year (*vmy*), the occupant age, and the crash environment of the crash year on injury severity.

[1] Let U be some additional covariate with outcomes u_l. If $\mathbb{P}(Y_N = y_i | X_N = x_j, Z_N = z_k, U_N = u_l) = \mathbb{P}(Y_I = y_i | X_I = x_j, Z_I = z_k, U_I = u_l)$ for all i, j, k, l, and $\mathbb{P}(U_N = U_l | X_N = x_j, Z_N = z_k) = \mathbb{P}(U_I = u_l | X_I = x_j, Z_I = z_k)$ for all j, k, l, then $\mathbb{P}(Y_N = y_i | X_N = x_j, Z_N = z_k) = \mathbb{P}(Y_I = y_i | X_I = x_j, Z_I = z_k)$ for all i, j, k.

Data Sets

National Data Set (DESTATIS)

The DESTATIS data set is the pool of all police-reported traffic crashes in all 16 German states. However, for data protection reasons, only data subsets or partial, i.e., resampled data sets with aggregated and pseudonymised variables, are available for independent research. To circumvent this limitation, the data subset used in this investigation consisted of multidimensional marginal distributions of user-defined variables (see Section III (Homogeneous Subgroups/Cut-off Points)) spanning the crash years 2010 to 2020 and focusing on the drivers of passenger cars.

In total, this data set contains information on approximately 4 million drivers.

In-depth Sample (GIDAS)

The GIDAS data set is a random sample of crashes with an injured party in two of Germany's metropolitan areas: Greater Dresden and Greater Hanover. Note that the metropolitan areas include the cities and the surrounding rural regions. The split between city areas and the rural regions was chosen to represent the split in Germany. The implemented on-spot, 50 % sampling is done in 6 *h* on/off shifts, alternating weekly, by a single team in each location. After completing a case, the most recent crash is addressed (first-in, last-out) until all crashes are sampled or the end of the shift is reached. Overall, this leads to a biased 30 % sample of crashes compared to DESTATIS [4][12].

The GIDAS sample is a non-identifiable sub-sample of the DESTATIS data set. The GIDAS data set from July 2021 comprises crashes between 2000-2020. For this analysis, the GIDAS sampling period is oriented on the DESTATIS data set; thus, only crashes between 2010-2020 are used. This subset contains information on about 20,000 crashes, of which about 15,500 are passenger vehicle crashes with roughly 30,000 occupants.

GIDAS variables beyond variables encoded in DESTATIS were used only for the injury severity evaluation. Note that this data set also contains information on passengers, specifically front-row passengers.

Injury Coding in the Data Sets

The P-Scale (injury severity documented by police officers)

In the DESTATIS data set, injury severity is categorised—by the police—into four levels: uninjured, slightly injured, severely injured, and fatal; these four levels of injury are denoted henceforth as {P0,P1,P2,P3}, respectively.

Uninjured is defined as no self-reported injuries and no need for medical assistance at the accident scene. Anybody who is either self-reporting an injury or is treated for a crash-induced medical condition is coded as at least slightly injured. To be coded as severely injured, the patient must be hospitalised for crash-related injuries. By international agreement, the fatal category is adjusted one month after the incident to include all crash-related fatalities within 30 days of the accident [13][7, Annex IV]. Thus, the P-scale is based only on treatment location, duration, and survival and not on injury severity evaluated by police officers. Consequently, this fact makes the P-scale more robust to misclassification than other policy-reported injury severity scales, such as the US KABCO scale and its related problems [14].

The GIDAS data set also includes the four-level police-reported injury severity.

S-Scale (injury severity using AIS-coding)

GIDAS also includes an AIS [15-16] coding of every individual injury [4]. In a previous project, all GIDAS injury information was reevaluated and recoded to the current AIS 2015 [17-18]. The AIS assesses the severity of individual injuries; these injury severities have to be aggregated to the patient level. For this injury severity aggregation, the MAIS of a patient and injury severity score (ISS) [19-20] are frequently employed. Therefore, an injury severity categorization using the well-established cut-off points of the AIS-based ISS is chosen for the evaluation and is denoted S-scale. At the cut-off points, the ISS gives a reasonable ordering; thus, this injury grouping can also be characterised as a scaling. It contains four levels of injury severity with the following groups:

S0 for uninjured (ISS ≤ 3),
S1 for light and moderate injuries (ISS $\in [4,8]$),
S2 for severe and life-threatening injuries with a high probability of survival (ISS $\in [9,15]$)
S3 for life-threatening injuries and fatalities (ISS ≥ 16)

Groups S2 and S3 are motivated by the established cut-offs of ISS ≥ 9 and ISS ≥ 16 [21-22] for severe and multiple injuries, respectively. Furthermore, the reliance on self-reporting in data sets such as GIDAS and DESTATIS for slight injuries, i.e., AIS1 injuries, results in an under-reporting of these injuries [23] and hinders differentiation between the MAIS0 and MAIS1 injured [24][4]. Therefore, both injury severity levels were used to represent S0.

To harmonise crash research, the European Commission issued a policy statement on road safety in which serious injuries are defined as those with a MAIS3 or higher, which approximately corresponds to ISS ≥ 9. Note that $S \in \{S2,S3\}$ refers to the European Commission's definition of serious injuries, i.e., MAIS3+. Furthermore, MAIS2+ can be expressed as $S \in \{S1,S2,S3\}$, and MAIS4+ is equivalent to $S \in \{S3\}$.

Homogeneous Subgroups/Cut-off Points

For *mapping of conditional probabilities* to work, it is necessary to classify the data into subgroups homogeneous with respect to the conditional probabilities across data sets. Note that the subgroups have to be defined identically for both data sets, further limiting possible cut-off points. Therefore, the pseudo-continuous data used in the analysis were categorised by cut-off points subject to objectively logical constraints, privacy constraints, and sub-sample size requirements.

Occupant Age

The occupant age is categorised into four groups: 18-44, 45-46, 65-74, and 75+ years of age (*yoa*). Since the focus is on drivers and the legal driving age in Germany is 18 or older, any group of occupants aged 17 or younger is not considered. Furthermore, the adult population (18-64 *yoa*)[25] is subcategorised into a young adult group (18-44 *yoa*)[26] and an older adult group (45-64 *yoa*).

The retiree group—starting at 65 *yoa*—is split at 75 *yoa* to consider the increase of fragility and frailty with age [27]. A stratification of the retiree group with finer granularity, i.e., into 5 *yoa* blocks, as suggested by [28], resulted in sample size problems.

Vehicle Model Year

Vehicle model year (*vmy*) is widely used as a proxy for crashworthiness [29]. Due to the unavailability of *vmy* in the data sets, the year of the first passenger car registration is used as a substitute. Furthermore, the year of first registration of the passenger car is only available for vehicles registered in Germany, about 95% of the sampled cars.

The current evaluation considers six categories: before 1980, 1980-1997, 1998-2001, 2002-2005, 2006-2011, and 2012-2020. Since the data set DESTATIS covers the crash years 2010-2020, the group *before 1980* contains vintage vehicles only and is not investigated in this analysis. Table AI gives the shares of each vehicle group for each period.

The cut-off points are motivated by the following arguments [30]:

1997/1998 Introduction of European New Car Assessment Programme (Euro NCAP) offset deformable Barrier (ODB) Test at 64 km/h rating [31].

2001/2002 compliance of new vehicle designs to Euro NCAP ODB.

2005/2006 compliance of all new vehicles to Euro NCAP ODB.

2011/2012 introduction of the Euro NCAP rigid pole side crash test [32].

Traffic Domains

Within the operating range of the restraint systems, motor vehicles provide very good protection to belted occupants. Furthermore, the injuries sustained despite this protection are overwhelmingly a result of blunt force trauma. The injury mechanism changes once the crash energy is above the design limit, i.e., the crumple zone and the forward excursion are used up.

Because of the lack of crash reconstruction data in the DESTATIS database, the speed limit at the crash site is used as a proxy for crash severity [33]. We found the speed limit at which the injury mechanism changes in a statistically significant way to be between 80 and 90 *km/h*, i.e., occurrence of sharp and penetrating traumata in a significant number of cases. Therefore, traffic domains are stratified by the speed limit (*SL*): *SL* $\leq$ 80 *km/h* and *SL* $\geq$ 90 *km/h* or Autobahn.

The change of injury mechanism has to be considered when assessing the percentage of drivers hospitalised for crash related injuries while MAIS3+ injured and surviving (see Section III (Mapping)).

Crash Year

The crash sample at hand encompasses crash years 2010 to 2020. The primary focus of our analysis is on the most recent vehicles, i.e., those with a market introduction in 2012 or later. While the first two crash years contain crashes before the market introduction of 2012 model year vehicles, the latter years have an increasing number of these newer cars (see the last column of Table AI).

The crash year was controlled to take care of changes in the crash environment, i.e., a change in injury severity of a given accident situation over time, such as progressively more restrictive speed limits and the installation of further crash barriers, but also weight increases in the crash opponents. Though these effects are minor year-to-year, they could be significant over the decade under investigation.

A three-year moving average was chosen to smooth the data, a trade-off between sample size and selection bias. When re-combining the crashes for information on the 2012–2020 crash years, the distributions are averaged across the three-year moving averages so that each crash year contributes equally.

Mapping of DESTATIS P-Coding to the S-Scale

The Link between P- and S-Scale

The DESTATIS data set does not contain any injury severity coding in S-scale or MAIS information. Hence, *mapping of conditional probabilities* supplements DESTATIS with injury severity coding in the S-scale. The variable P refers to the injury distribution in the P-scale, and the variable S refers to the S-scale injury distribution. Both variables have four possible outcomes. We have denoted the outcomes P0,P1,P2,P3 and 0,1,2,3, respectively. Additionally, let $\mathbb{P}(P = \mathrm{P}j) = x_j$ and $\mathbb{P}(S = \mathrm{S}j) = y_j$ for $j = 0, \dots, 3$. In the DESTATIS data set, the probabilities x_0, x_1, x_2,and x_3 were given, and we used the *mapping of conditional probabilities* approach to determine the probabilities y_0, y_1, y_2,and y_3 for the DESTATIS data set.

We follow [24] and do not distinguish between P0 and P1. The reason for this is twofold. First, the focus of this application is on severely injured persons, which are encoded primarily by P2 or P3. Second, [24] argue that the relying on self-reporting for slight injuries hinders a differentiation between the MAIS0 and MAIS1 injured. Consequently, the coding of P0 and P1 is not considered trustworthy. Hence, the variable P is reduced to 3 outcomes, and we have denoted its outcomes as $\mathrm{P0} + \mathrm{P1}$,P2,and P3. Thus, for the variables S and P, we got $4 \times 3 = 12$ conditional probabilities: $\mathbb{P}(S = \mathrm{S}i | P = \mathrm{P0} + \mathrm{P1}), \mathbb{P}(S = \mathrm{S}i | P = \mathrm{P2}), \mathbb{P}(S = \mathrm{S}i | P = \mathrm{P3}), i = 0,1,2,3$. These conditional probabilities can be estimated using the GIDAS data set. Before doing so, note the constraint $\sum_{i=0}^{3} \mathbb{P}\,(S = \mathrm{S}i | P = \omega) = 1, \omega \in \{\mathrm{P0} + \mathrm{P1}, \mathrm{P2}, \mathrm{P3}\}$ reduced the unconstrained parameters to $9 = (4 - 1) \times 3$. Additionally, some conditional probabilities did not need to be computed since they can be set to zero. To elaborate, we set $\mathbb{P}(S \geq \mathrm{S1} | P = \mathrm{P0} + \mathrm{P1}) = 0$ since almost all AIS2 injuries would warrant a hospital stay, at least in Germany. Moreover, since the probability of dying from an AIS2 injury is negligible, i.e., a coding error in either S or P is much more likely than the patient dying from one or two AIS2 injuries, we set $\mathbb{P}(S \leq \mathrm{S1} | P = \mathrm{P3}) = 0$.

Overall, we obtained the following

$$
\begin{aligned}
&\mathbb{P}(S = \mathrm{S0} | P = \mathrm{P0} + \mathrm{P1}) = 1, && \mathbb{P}(S = \mathrm{S0} | P = \mathrm{P2}) = \alpha_0, && \mathbb{P}(S = \mathrm{S0} | P = \mathrm{P3}) = 0\\
&\mathbb{P}(S = \mathrm{S1} | P = \mathrm{P0} + \mathrm{P1}) = 0, && \mathbb{P}(S = \mathrm{S1} | P = \mathrm{P2}) = \alpha_1, && \mathbb{P}(S = \mathrm{S1} | P = \mathrm{P3}) = 0,\\
&\mathbb{P}(S = \mathrm{S2} | P = \mathrm{P0} + \mathrm{P1}) = 0, && \mathbb{P}(S = \mathrm{S2} | P = \mathrm{P2}) = \alpha_2, && \mathbb{P}(S = \mathrm{S2} | P = \mathrm{P3}) = 1 - \gamma,\\
&\mathbb{P}(S = \mathrm{S3} | P = \mathrm{P0} + \mathrm{P1}) = 0, && \mathbb{P}(S = \mathrm{S3} | P = \mathrm{P2}) = \alpha_3, && \mathbb{P}(S = \mathrm{S3} | P = \mathrm{P3}) = \gamma,
\end{aligned}
$$

where $\alpha_3 = 1 - \alpha_2 - \alpha_1 - \alpha_0, \alpha_0, \alpha_1, \alpha_2$, and $\gamma \in [0,1]$.

The *mapping of conditional probabilities* approach transfers these conditional probabilities from GIDAS to DESTATIS. In the context of this application, the population level refers to drivers of passenger vehicles. Recall that the underlying assumption is that these conditional probabilities coincide. Most likely, this assumption does not hold for the population level but should hold for the subgroups; see the discussion of Equation 2 for details.

The α parameters give the shares of S0 to S3 for survivors of a traffic crash who were hospitalised for crash related causes. The parameter γ links fatality to injury severity, i.e., γ provides the share of the S2 and S3 injured for the non-survivors of a traffic crash. As a pure injury metric, technical crash severity should not affect γ. Since fragility affects the survival probability, γ is affected by age. The injury mechanism and the medical treatment policy of traffic crashes affect the α parameters. Since both DESTATIS and GIDAS are German data sets, the medical treatment policy for traffic crashes can be seen as similar. If there is no intrusion into the occupant compartment, the injury mechanism of traffic crashes is mainly blunt force trauma. Frailty, fragility, and medical treatment policy differences in occupant age make α age-dependent. Thus, to take into account these dependencies and differences in the age distribution between DESTATIS and GIDAS, the parameters α and γ are stratified by age.

Additionally, the α parameters are stratified into two groups by the speed limit (*SL*) at the crash location. These two groups can be seen as a proxy for the probability of significant intrusion into the vehicle's passenger compartment. The vehicle model year affects the stiffness and restraint system, affecting the likelihood of intrusion and the injury mechanism. As seen in Table AI, the vehicle model year in traffic crashes changes relatively rapidly. If the α and γ parameters were not stratified by vehicle model year, this temporal instability could be transferred to these parameters, see [9] for a discussion about the issues of temporal unstable parameters. Considering the stratifications mentioned above, the parameters α and γ are temporally stable from 2010 to 2020 since the crash year is statistically insignificant, see Appendix (Temporal stability of the α parameter).

Below, we list such subgroups for each parameter.

α_i: age of occupant [*yoa*] ∈ $\{18-44, 45-64, 65-74, 75+\}$ *yoa,*
speed limit (*SL*) at crash location [*km/h*]: *SL* ∈ $\{5-80, 90+\}$ *km/h*,
vehicle model year (*vmy*) ∈ $\{1980-1997, 1998-2001, 2002-2005, 2006-2011, 2012-2020\}$.

γ: age of occupant [*yoa*] ∈ $\{18-44, 45-64, 65-74, 75+\}$ *yoa*.

Note that, as stated in Section II, stratification of the conditional probabilities is necessary only in dimensions in which the distribution of the stratification variable differs among the data sets.

Hence, for each subgroup, given the parameters $\alpha_i, i \in \{0,1,2\}$ and γ the P-scale can be mapped to the S-scale. This defines the following mapping (on-subgroup level)

$\tilde{S}(x)$: $\{\text{P0}: x_0, \text{P1}: x_1, \text{P2}: x_2, \text{P3}: x_3\} \to \{\text{S0}: y_0, \text{S1}: y_1, \text{S2}: y_2, \text{S3}: y_3\}$,

$$\tilde{S}(x) = \begin{cases} \text{S0}: y_0 & = x_0 + x_1 + \alpha_0 x_2, \\ \text{S1}: y_1 & = \alpha_1 x_2, \\ \text{S2}: y_2 & = \alpha_2 x_2 + (1-\gamma) x_3, \\ \text{S3}: y_3 & = (1 - \alpha_0 - \alpha_1 - \alpha_2) x_2 + \gamma x_3 = \alpha_3 x_2 + \gamma x_3. \end{cases} \tag{6}$$

Estimation of the α_i, and γ Parameters

The parameters $\alpha_i, i \in \{1,2,3\}$, and γ need to be estimated for each subgroup using the in-depth data set GIDAS. For each α_i parameter, there are $4 \times 2 \times 5 = 40$ subgroups ([age of occupant] + [speed limit] + [vehicle model year]). A logistic regression is used to address the resulting sample size problems. To elaborate, let $\mathbb{1}$ be the indicator function. The following logistic model is used

$$\alpha_i = \frac{1}{1 + \exp\left(-\left(\beta_0 + \beta_1 \,\text{yoa} + \beta_2 \,\text{yoa}^2 + \beta_3 \mathbf{1}(\text{SL} \in \{90+\}) + \sum_{i=1}^{5} \beta_{i+4} \mathbf{1}(\text{vmy} \in \text{vmy}_\text{i})\right)\right)}.$$

That means the occupant's age is used numerically (as a polynomial of order 2 to take the drastically increasing fragility and frailty of older people into account) [27]. In contrast, the speed limit at the crash location and vehicle model year are used categorically in a one-hot encoding. Thus, the number of parameters for each α_i is reduced to 9.

The parameter γ is estimated model-free. Since GIDAS contains information on the driver and all passengers, all parameters are estimated in GIDAS using the information on all front-row occupants. Hence, it is assumed that the injury mechanism is similar for drivers and other front-row occupants. The γ parameters are conditional probabilities and can be estimated by relative frequencies in the GIDAS data set. Relative frequencies are also consistent estimators in the independent and identically distributed (i.i.d.) setting [34]. Furthermore, bootstrap-based statistical inference for these parameters and the logistic regression can be obtained [35-36].

The mapping (Equation 6) applied to the DESTATIS data set has as inputs the injury distribution in levels {P0,P1,P2,P3}, which are determined using the DESTATIS data set. We treat the data set DESTATIS as a random i.i.d. sample for each crash year. Thus, the relative frequencies of {P0,P1,P2,P3} are considered estimates of the unknown injury severity risk. Therefore, the obtained relative frequencies in the S-scale are affected by the statistical uncertainty of both data sets. The statistical inference is obtained by a joint bootstrap-based inference of both data sets using an i.i.d. resampling scheme to consider this uncertainty in our statistical analysis. The obtained 95 %-confidence intervals are based on $B = 10\,000$ bootstrap repetitions.

Thus, the evaluation is based on the following assumptions:

1. GIDAS is a subsample of DESTATIS, and both data sets coincide in encoding practice for variables present in both data sets, especially P-scale coding, vehicle model year, occupant age, and the speed limit at the crash location.
2. The conditional probabilities of the S-scale given P-scale coincide in both data sets for the clusters defined by vehicle model year, occupant age, and the speed limit at the crash location.
3. The clusters themself are defined homogeneously, i.e., the frequency distributions in both data sets coincide, or the conditional probabilities of the öS-scale given P-scale are homogenous within each cluster.
4. A logistic model describes the conditional probability of the S-scale given P2 across the defined clusters.

IV. RESULTS

The estimated $\alpha_i, i = 0,1,2,3$, and γ-parameters—stratified by age group and vehicle model year—calculated from the GIDAS data set are given in Table AIII in the Appendix.

Timeline Trends on the S-Scale

The DESTATIS data subset contained crashes from the years 2010-2020. Though for some results, the data set has been cropped to 2012-2020 as the primary focus was on the current traffic situation of modern vehicles, i.e., those with a vehicle model year 2012-2020.

Crash Years

The S-scale injury severity distribution for each group of the three-year moving average and the aggregated injury severity information for crash years 2012-2020 are given in Table AII. This information is visualised in Figure A6 in Appendix (Injury Severity by crash-year and age group). All $S0 \rightarrow S1$ and $S1 \rightarrow S2$ transitions between the crash years are statistically significant.

The tabulation of the S-scale trends for the crash year groups stratified by age groups is given in Table AIV and Figure A7 in Appendix (Injury Severity by crash-year and age group).

Age Groups

The shift in injury severity with an increase in occupant age—while focusing on crash years 2012-2020 (weighted) and vehicle model years 2012-2020—is visualised in Figure A3.

At the $S3$ and $S2$ levels, the relative injury severity frequencies increase with age, while at the $S1$ level, the 18–44 *yoa* and 45–64 *yoa* groups are nearly indifferent, with a marked increase for the 65+ groups.

Vehicle Model Year

The reduction in the relative frequency of injury severity with newer vehicles is shown in Figure A4. The shift in injury severity is depicted for vehicle model year groups for the crash years 2012-2020. The shift is more substantial for lower S-scores than for higher S-scores, and all $S1 \rightarrow S0$ shifts are statistically significant on the 95%-level.

Age Group and Vehicle Model Year

Age group stratified relative injury frequencies in crashed vehicles from the *pre-NCAP* era (1980–1997 *vmy*) are compared to the current fleet (2012-2020 *vmy*) in Figure A5.

Compared to the *pre-NCAP* vehicles, the injury severity situation in modern cars improved in all age groups. There are even some shifts from $S3 \rightarrow S1$ in the two younger age groups. Furthermore, there are some significant injury severity shifts in the older age groups, such as the 50 % $S3$ reduction in the 75+ *yoa* group. Thus, again, the shift is more substantial for lower S-scores than higher ones.

The more detailed plots with all the intermediate vehicle model year groups and stratification by age groups can be found in Figure A8 in in Appendix (Injury Severity by crash-year and age group). For vehicle model year groups $[98{,}01]$, $[02{,}05]$, and $[06{,}11]$, there are insignificant changes for all age groups in both directions between $S1$ and $S2$ and $S2$ and $S3$. The injury severity reduction is relatively more substantial in the younger age groups on all injury severity levels. In the two older age groups, there is only a significant change in $S3 \rightarrow S2$ for vehicle groups $[80{,}97] \rightarrow [98{,}01]$.

V. DISCUSSION

MAIS3+ *Mapping to the National Level*

Despite the police not measuring MAIS3+ in their data, distributions of their P-scale of crash severity can be used to compute distribution information on MAIS3+ (S2) injury severity. The method *mapping of conditional probabilities* used the GIDAS data set as the source for this information, as all those accidents are coded using P- and S-scales simultaneously.

Application to Timelines

The obvious application of new injury severity assessment in crash data is timelines: injury severity over crash year, vehicle model year, or occupant age.

Modern vehicles are safer than older ones; this result is visualised in Figure A4. Thus, with the increase of more modern vehicles on the road (see Table AI), the relative frequency of MAIS3+ injuries is expected to decrease. Furthermore, not only are *S2* and *S2+* decreasing with vehicle model year, but so are *S1*, *S1+*, and *S3*.

When comparing the real-world crash performance of pre-NCAP vehicles to post-NCAP vehicles, some significant injury severity shifts become apparent (Figure A5). For the most severe injury levels (S2 and S3), more than half of the injuries are shifted down by one severity level, i.e., *S3→S2* and *S2→S1*. This shift also holds for the *S1* level, except for the oldest demographic. Furthermore, there are small but statistically significant injury severity shifts from *S3→S1* in the two younger age groups. The proportion of the change in injury severity is most prominent for the youngest demographic and decreases significantly with occupant age.

A more detailed look at the safety improvements stratified by age group is shown in Figure A8. Major shifts from *S3→S2* are present only for *vmy* [80,97]→[97,01]. Afterwards, only minor (most even statistically insignificant) changes are visible. Major shifts from *S2→S1* are seen in all age groups for vmy [80,97]→[97,01] and [06,11]→[12,20]. Again, the technological advancements benefitted the younger demographic more than the older demographic; this is especially obvious at the *S1+* level but holds for all injury severity levels.

Thus, the most potential for injury severity reduction can be found in the oldest demographic. However, care must be taken not to reduce the injury severity for the old age groups at the expense of the younger age groups [37].

Vehicle Model Years 2012-2020, Only

The timelines show that the latest vehicle models provide the best occupant protection, independent of the occupant's age (see Figure A4 and Figure A8). Therefore, this group (2012-2020 *vmy*) was chosen for further investigation.

The injury severity level sustained by a vehicle occupant is age-dependent: The older the occupants, the more severe the injury severity level; the increase is most pronounced for the two most senior age groups, see Figure A3. This age-dependency can be directly attributed to increased fragility and frailty with age [27]. Likewise, this age-dependency of the injury severity level also holds when focusing on the vehicle model years 2012-2020 only (see Figure A5).

Within the crash years 2012-2020, occupants in vehicles with a vehicle model years 2012-2020 fare worse the more current the crash year, i.e., the cash environment is becoming more and more aggressive for the modern cars. This deterioration in crash performance holds for the injury severity levels $S1$, $S2$, and $S3$ but also for P2 and P3, whose relative frequencies are increasing (see Table AII and Figure A6). This trend also holds when this population is stratified by age group (see Table AV).

As crash opponents, the increased number of in newer vehicles (see Table AI) seems to increase injury severity. Even so, modern vehicles have a good crash performance as long as they crash against older vehicles (see Table AII and Figure A6). Occupants in newer vehicles do not seem to profit as much from increased safety when crashing into more recent vehicles.

To investigate the reason for this more aggressive crash environment is out of scope of this paper but clearly an avenue for further research. One of the reasons could be the ever-increasing vehicle mass [38-39] and the consecutive ever-stiffer front ends due to the crash-test focus on self-protection, i.e., test speeds independent of vehicle mass [30]. Another reason could be an increase in aggressive driving behaviour [40].

Limitations

DESTATIS *Sampling Limits*

The DESTATIS database consists of police-recorded crashes only. Thus, there are some *property damage only* crashes missing. In addition, minor injuries, e.g., hematoma or superficial lacerations, will not be reported for these cases, as there is an under-reporting at the lower end of the injury severity spectrum.

All the variables in the DESTATIS database are recorded by the police officers. However, some information of little interest to the police investigation, e.g., seat belt status and seating position within the vehicle other than the driver, have quality issues or are not documented at all, e.g., height and weight of the occupant. Furthermore, some pieces of information on the crash are the result of the crash reconstruction, e.g., the crash sequence, which is neither done for every case nor are the results available at the time the police officer has to file the crash information with the statistics office (30 days post-crash). Therefore, all these variables with quality issues cannot be used in a re-coding process like the one described in this paper.

The current analysis used pre-computed, multi-dimensional marginal distributions computed with access to the 100 % DESTATIS sample. Increases in dimensionality are limited by data protection requirements, i.e., the ability to re-identify individual crashes.

GIDAS *Sample Size Limits*

The conditional probabilities, i.e., the α_i- and γ-parameters, are solely based on the GIDAS data set. Note that mapping these conditional probabilities to the DESTATIS data set rests on assumptions. An increase in the GIDAS sample size would significantly reduce the confidence intervals of the translation parameters (see Table AIII). Additionally, this would allow us to estimate the parameters model-free.

VI. CONCLUSION

Instead of projecting a small, in-depth sample to a larger level, a novel method was presented to map the marginal distributions of a large data set to a finer granularity coding using an in-depth data source to estimate required conditional distributions. Using marginal distributions instead of the entire data set addresses data protection concerns.

As an application of the method, the German national motor vehicle crash database (DESTATIS) was enhanced by distribution information on the MAIS2+*,* MAIS3+*,* and MAIS4+ injury severity levels from GIDAS.

VII. REFERENCES

[1] European Commission. (2019) EU Road Safety Policy Framework 2021-2030 – Next Steps towards 'Vision Zero'.

[2] Zhang, F., Chen, C-L. (2013) National Highway Traffic Safety Administration (NHTSA), NASS-CDS: Sample Design and Weights (DOT HS 811 807), Washington, D.C.

[3] Hansen, M H., Hurwitz, W N., Madow, W G. (1953) Sample Survey Methods and Theory - Volume I (Methods and Applications), John Wiley & Sons, New York.

[4] Otte, D., Krettek, C., Brunner, H., Zwipp, H. (2003) Scientific Approach and Methodology of a New In-depth Investigation Study in Germany so called GIDAS. Enhanced Safety Vehicles (ESV), 2003, Nagano, Japan.

[5] Hautzinger, H., Pfeiffer, M., Schmidt, J. (2006) Bundesanstalt für Straßenwesen (BASt), F59, Hochrechnung von Daten aus Erhebungen am Unfallort [Data Expansion of Accident Data from In-depth Accident Surveys], Bergisch Gladbach, Germany.

[6] Kreiss, J-P., Feng, G. et al. (2015) Extrapolation of GIDAS Accident Data to Europe, Enhanced Safety of Vehicles (ESV), 2015, Gothenburg, Sweden.

[7] European Commission. (2013) On the implementation of objective 6 of the European Commission's policy orientations on road safety 2011-2020 – First milestone towards an injury strategy, p.4.1. SWD(2013) 94 final.

[8] Behnood, A., Mannering, F. (2015) The temporal stability of factors affecting driver-injury severities in single-vehicle crashes: Some empirical evidence. Analytic Methods in Accident Research, **8**(1): p.7–32.

[9] Mannering, F. (2018) Temporal Instability and the Analysis of Highway Accident Data. Analytic Methods in Accident Research, **17**(1): p.1–13.

[10] European Commission. (2001) European transport policy for 2010: Time to decide}. COM(2001) 370 final.

[11] Lefering, R. (2009), Bundesanstalt für Straßenwesen (BASt), M200, Entwicklung der Anzahl Schwerstverletzter infolge von Straßenverkehrsunfällen in Deutschland [Development of the Number of Persons Critically Injured as a Result of Road Traffic Accidents in Germany], Bergisch Gladbach, Germany.

[12] Hautzinger, H., Pfeiffer, M., Schmidt, J. (2004) Expansion of GIDAS Sample Data to the Regional Level. Expert Symposium on Accident Research (ESAR), 2004, Hanover, Germany,

[13] Statistisches Bundesamt. (2022) DESTATIS, Sonderauswertung Straßenverkehrsunfälle 2010-2020: Beteiligte Pkw-Fahrer an Unfällen mit Personenschaden nach Erstzulassungsjahr des Pkws, Alter und Verletzungsschwere des Fahrers und Geschwindigkeitsbegrenzung am Unfallort [Special Evaluation of Road Traffic Crashes 2010-2020: Vehicle drivers involved in crashes with injuries stratified by year of vehicle registration, driver age and injury severity, and speed limit at the crash site], Wiesbaden, Germany.

[14] Farmer, CM. (2003) Reliability of Police-Reported Information for Determining Crash and Injury Severity. Traffic Injury Prevention, **4**(1): pp. 38–44.

[15] Gennarelli, T A. (1998) The Abbreviated Injury Scale 1990 - Update 98. Association for the Advancement of Automotive Medicine (AAAM), Barrington, IL.

[16] Gennarelli, T A., Wodzin, E. (2008) The Abbreviated Injury Scale 2005 - Update 2008. Association for the

Advancement of Automotive Medicine (AAAM), Barrington, IL.

[17] Gillich, P. (2015) The Abbreviated Injury Scale 2015 Revision. Association for the Advancement of Automotive Medicine (AAAM), Barrington, IL.

[18] Unger, T., Liers, H., Schuster, R., Kleber, C. (2019) Verband der Automobilindustrie (VDA), Entwicklung der Verletzungsschwere bei Verkehrsunfällen in Deutschland im Kontext verschiedener AIS-Revisionen [The Influence of AIS-Revisions on the Temporal Shift in Injury Severity in Road Traffic Crashes in Germany.] (FAT-Schriftenreihe 327), Berlin, Germany.

[19] Baker, S P., O'Neill, B. (1976) The Injury Severity Score: An Update. The Journal of Trauma, **16**(11): p. 882–885.

[20] Baker, S P., O'Neill, B., Haddon, W., Long, W. (1974) The Injury Severity Score: A Method for Describing Patients with Multiple Injuries and Evaluating Emergency Care. The Journal of Trauma, **14**(3): p.187–196.

[21] DGU Guideline Committee of the German Registered Society for Trauma Surgery. (2001) Recommended Guidelines for Diagnostics and Therapy in Trauma Surgery. European Journal of Trauma, **27**(3): p.137–150.

[22] Pape, H-C., Zelle, B. et al (2006) Evaluation and Outcome of Patients after Polytrauma – Can Patients be Recruited for Long-term Follow-up? Injury, **37**(12): p.1197–1203.

[23] Andricevic, N., Junge, M., Krampe, J. (2018) Injury risk functions for frontal oblique collisions. Traffic Injury Prevention, **19**(5): p.518–522.

[24] Krampe, J., Junge, M. (2021) Deriving functional safety (ISO 26262) S-parameters for vulnerable road users from national crash data. Accident Analysis & Prevention, **150**(1): p.1–9.

[25] Morris, A., Welsh, R., Frampton, R., Charlton, J., Fildes, B. (2002) An Overview of Requirements for the Crash Protection of Older Drivers. Annals of Advances in Automotive Medicine/Annual Scientific Conference, **46**: p.1–16.

[26] Bahouth, G., Digges, K., Schulman, C. (2012) Influence of injury risk thresholds on the performance of an algorithm to predict crashes with serious injuries. Annals of Advances in Automotive Medicine/Annual Scientific Conference, **56**: p.223–230.

[27] Kent, R., Trowbridge, M., Lopez-Valdes, F J., Heredero Ordoyo, R., Segui-Gomez, M. (2009) How Many People Are Injured and Killed as a Result of Aging? Frailty, Fragility, and the Elderly Risk-Exposure Trade-off Assessed via a Risk Saturation Model. Annals of Advances in Automotive Medicine/Annual Scientific Conference, **53**: p.41–50.

[28] Collard, R M., Boter, H., Schoevers, R A., Oude Voshaar, R C. (2012) Prevalence of Frailty in Community-Dwelling Older Persons: A Systematic Review. Journal of the American Geriatric Society, **60**(8): p.1487–1492.

[29] Ryb, G E., Dischinger, P C., McGwin, G., Griffin, R L (2011) Crash-related Mortality and Model year: Are Newer Vehicles Safer? Annals of Advances in Automotive Medicine/Annual Scientific Conference, **55**: p.113–121.

[30] van Ratingen, M., Williams, A. et al. (2016) The European New Car Assessment Programme: A Historical Review. Chinese Journal of Traumatology, **19**(2): p. 63–69.

[31] Euro NCAP. (2018) European New Car Assessment Programme, Offset Deformable Barrier Frontal Impact Testing Protocol, Version 7.1.3., Antwerp, Belgium.

[32] Euro NCAP. (2019) European New Car Assessment Programme, Oblique Pole Side Impact Testing Protocol, Version 7.1.1., Antwerp, Belgium.

[33] Krampe, J., Junge, M. (2020) Injury Severity for Hazard & Risk Analyses: Calculation of ISO 26262 S-parameter Values from Real-World Crash Data. Accident Analysis & Prevention **136**(1): p.1–8.

[34] Mood, A M., Graybill, F A., Boes, D C. (1974) Introduction to the Theory of Statistics, Chpt. VI.3. McGraw-Hill, New York.

[35] Efron, B., Tibshirani, R J. (1993) An Introduction to the Bootstrap, Chpts. 5, 21. Chapman & Hall/CRC, Boca Raton.

[36] Davison, A C., Hinkley, D V. (1997) Bootstrap Methods and Their Application, Sec. 7.2. Cambridge University Press, Cambridge, U.K.

[37] Krampe, J., Junge, M. (2019) Population-based assessment of a vehicle fleet with seat belts providing lower shoulder belt forces than today. Traffic Injury Prevention, **20**(3): p.320–324.

[38] Kahane, C J. (2003) National Highway Traffic Safety Administration (NHTSA), Vehicle Weight, Fatality Risk and Crash Compatibility of Model Year 1991–99 – Passenger Cars and Light Trucks (DOT HS 809 662), Washington, D.C.

[39] Highway Loss Data Institute (HLDI). (2011) Injury Odds and Vehicle Weight - Comparison of Hybrids and Conventional Counterparts}. Highway Loss Data Institute Bulletin. **28**(10): p.1–10.

[40] Islam, M., Mannering, F. (2020) A temporal analysis of driver-injury severities in crashes involving aggressive and non-aggressive driving. Analytic Methods in Accident Research, 27: p.1–16.

[41] DESTATIS (2018) Statistisches Bundesamt (DESTATIS), Fachserie 8 Reihe 7: Verkehr – Verkehrsunfälle [traffic – crashes], pp. 87, 99, Wiesbaden, Germany.

[42] Institut für angewandte Sozialwissenschaft (infas). (2017) Bundesministerium für Verkehr und digitale Infrastruktur, Referat G 13, Mobilität in Deutschland – Tabellarische Grundauswertung [Mobility in Germany – Basic Tables], Tab. B A15, Berlin, Germany.

[43] IRTAD. (2014) OECD/ITF, Road Safety - Annual Report 2014, Tab 6, Paris, France.

VIII. APPENDIX

Tables

Crash year	year of first vehicle registration							
	NA	–79]	[80,97]	[98,01]	[02,05]	[06,11]	[12,20]	n
2010	0.055	0.001	0.213	0.218	0.220	0.294	0.000	354919
2011	0.057	0.001	0.181	0.206	0.212	0.344	0.000	370632
2012	0.056	0.001	0.150	0.193	0.205	0.362	0.034	367055
2013	0.058	0.001	0.119	0.175	0.198	0.351	0.099	359808
2014	0.096	0.001	0.092	0.152	0.181	0.325	0.153	371095
2015	0.057	0.001	0.077	0.142	0.179	0.326	0.218	378156
2016	0.058	0.001	0.063	0.125	0.168	0.313	0.273	381354
2017	0.057	0.001	0.050	0.108	0.154	0.300	0.330	372144
2018	0.055	0.001	0.039	0.091	0.140	0.285	0.390	369050
2019	0.055	0.001	0.031	0.075	0.126	0.271	0.441	357327
2020	0.073	0.001	0.026	0.063	0.116	0.252	0.470	286079
n	243453	2550	379506	563378	688856	1239831	850045	3967619

Table AI
Shares of vehicle to market introduction groups for several time periods. (NA: not available).

	crash year group			
	2012-2014	2013-2015	2014-2016	2015-2017
S0	97.43(97.20,97.66)	97.36(97.12,97.59)	97.24(97.00,97.48)	97.11(96.85,97.36)
S1	1.66(1.43,1.89)	1.70(1.47,1.94)	1.78(1.54,2.02)	1.87(1.62,2.13)
S2	0.49(0.35,0.64)	0.50(0.36,0.66)	0.53(0.37,0.69)	0.55(0.39,0.73)
S3	0.42(0.32,0.54)	0.43(0.33,0.55)	0.45(0.34,0.58)	0.47(0.35,0.60)
S1+	2.57(2.34,2.80)	2.64(2.41,2.88)	2.76(2.52,3.00)	2.89(2.64,3.15)
S2+	0.91(0.74,1.09)	0.94(0.76,1.12)	0.98(0.80,1.17)	1.02(0.83,1.22)
P2	3.99(3.91,4.08)	4.10(4.04,4.17)	4.28(4.22,4.33)	4.50(4.45,4.55)
P3	0.18(0.16,0.20)	0.18(0.17,0.20)	0.19(0.18,0.20)	0.19(0.18,0.20)
	crash year group			
	2016-2018	2017-2019	2018-2020	2012-2020 (weighted)
S0	97.00(96.74,97.27)	96.90(96.63,97.17)	96.85(96.58,97.13)	97.13(96.88,97.38)
S1	1.94(1.68,2.21)	2.01(1.74,2.29)	2.04(1.76,2.32)	1.86(1.61,2.11)
S2	0.57(0.41,0.75)	0.59(0.42,0.78)	0.60(0.43,0.79)	0.55(0.39,0.72)
S3	0.48(0.36,0.61)	0.49(0.37,0.63)	0.51(0.38,0.65)	0.47(0.35,0.59)
S1+	3.00(2.73,3.26)	3.10(2.83,3.37)	3.15(2.87,3.42)	2.87(2.62,3.12)
S2+	1.05(0.86,1.26)	1.09(0.88,1.31)	1.11(0.90,1.33)	1.01(0.83,1.22)
P2	4.68(4.63,4.73)	4.84(4.80,4.89)	4.91(4.86,4.95)	4.47(4.44,4.51)
P3	0.19(0.18,0.20)	0.20(0.19,0.21)	0.21(0.20,0.22)	0.19(0.18,0.20)

Table AII
S-distribution for different crash year groups, vehicle model year group 2012-2020.

	occupant age:		$[18,44]$	$[45,64]$	$[65,74]$	$[75+]$	$[18,75+]$
vmy $\in [80,97]$	$SL \leq 80\,km/h$	α_0	0.34(0.30,0.38)	0.29(0.25,0.34)	0.30(0.25,0.35)	0.31(0.25,0.38)	0.32(0.28,0.35)
		α_1	0.44(0.40,0.48)	0.46(0.41,0.50)	0.44(0.39,0.49)	0.41(0.35,0.48)	0.44(0.40,0.48)
		α_2	0.15(0.12,0.18)	0.17(0.14,0.21)	0.18(0.15,0.22)	0.19(0.14,0.24)	0.17(0.14,0.19)
		α_3	0.07(0.05,0.09)	0.08(0.05,0.10)	0.08(0.06,0.11)	0.08(0.05,0.12)	0.07(0.05,0.09)
	$SL \geq 90\,km/h$	α_0	0.33(0.28,0.37)	0.27(0.22,0.33)	0.27(0.22,0.33)	0.29(0.21,0.36)	0.31(0.26,0.35)
		α_1	0.40(0.36,0.44)	0.41(0.36,0.47)	0.40(0.34,0.45)	0.37(0.31,0.44)	0.40(0.36,0.44)
		α_2	0.20(0.16,0.24)	0.22(0.18,0.27)	0.23(0.18,0.29)	0.24(0.18,0.31)	0.21(0.17,0.25)
		α_3	0.08(0.05,0.10)	0.09(0.06,0.12)	0.09(0.06,0.13)	0.10(0.06,0.14)	0.08(0.06,0.11)
vmy $\in [98,01]$	$SL \leq 80\,km/h$	α_0	0.39(0.29,0.48)	0.35(0.25,0.45)	0.35(0.25,0.45)	0.37(0.26,0.48)	0.37(0.27,0.46)
		α_1	0.43(0.33,0.54)	0.44(0.34,0.55)	0.43(0.33,0.53)	0.40(0.30,0.52)	0.43(0.33,0.53)
		α_2	0.17(0.10,0.24)	0.19(0.11,0.27)	0.20(0.12,0.29)	0.20(0.12,0.30)	0.18(0.11,0.26)
		α_3	0.02(0.00,0.05)	0.02(0.00,0.06)	0.02(0.00,0.06)	0.02(0.00,0.06)	0.02(0.00,0.05)
	$SL \geq 90\,km/h$	α_0	0.38(0.28,0.48)	0.33(0.23,0.44)	0.33(0.22,0.44)	0.35(0.23,0.47)	0.36(0.26,0.46)
		α_1	0.39(0.29,0.49)	0.40(0.30,0.51)	0.39(0.28,0.50)	0.36(0.26,0.48)	0.39(0.29,0.50)
		α_2	0.21(0.13,0.31)	0.24(0.14,0.35)	0.25(0.15,0.36)	0.26(0.15,0.38)	0.23(0.14,0.33)
		α_3	0.02(0.00,0.05)	0.02(0.00,0.07)	0.03(0.00,0.07)	0.03(0.00,0.07)	0.02(0.00,0.06)
vmy $\in [02,05]$	$SL \leq 80\,km/h$	α_0	0.37(0.30,0.44)	0.33(0.26,0.40)	0.34(0.27,0.41)	0.36(0.28,0.43)	0.35(0.29,0.42)
		α_1	0.46(0.39,0.53)	0.47(0.40,0.55)	0.46(0.38,0.53)	0.43(0.35,0.51)	0.46(0.39,0.53)
		α_2	0.11(0.07,0.16)	0.13(0.09,0.18)	0.14(0.09,0.19)	0.14(0.09,0.20)	0.12(0.08,0.17)
		α_3	0.06(0.03,0.09)	0.07(0.04,0.11)	0.07(0.04,0.11)	0.07(0.03,0.12)	0.06(0.03,0.10)
	$SL \geq 90\,km/h$	α_0	0.37(0.30,0.44)	0.32(0.25,0.40)	0.33(0.25,0.40)	0.34(0.26,0.43)	0.35(0.28,0.42)
		α_1	0.42(0.34,0.49)	0.43(0.36,0.51)	0.41(0.34,0.49)	0.39(0.31,0.48)	0.42(0.35,0.49)
		α_2	0.15(0.10,0.21)	0.17(0.11,0.24)	0.18(0.12,0.24)	0.18(0.12,0.26)	0.16(0.11,0.22)
		α_3	0.07(0.03,0.11)	0.08(0.04,0.13)	0.08(0.04,0.13)	0.08(0.04,0.14)	0.07(0.04,0.11)
vmy $\in [06,11]$	$SL \leq 80\,km/h$	α_0	0.38(0.33,0.43)	0.33(0.29,0.38)	0.34(0.29,0.39)	0.36(0.29,0.42)	0.36(0.32,0.40)
		α_1	0.41(0.36,0.47)	0.43(0.38,0.48)	0.41(0.36,0.46)	0.39(0.32,0.45)	0.42(0.37,0.46)
		α_2	0.14(0.11,0.18)	0.17(0.13,0.20)	0.17(0.14,0.21)	0.18(0.13,0.23)	0.16(0.13,0.19)
		α_3	0.06(0.04,0.09)	0.07(0.05,0.10)	0.07(0.05,0.10)	0.08(0.04,0.12)	0.07(0.05,0.09)
	$SL \geq 90\,km/h$	α_0	0.37(0.31,0.42)	0.32(0.26,0.37)	0.32(0.26,0.38)	0.33(0.26,0.41)	0.35(0.30,0.40)
		α_1	0.37(0.32,0.43)	0.39(0.33,0.44)	0.37(0.32,0.43)	0.35(0.28,0.42)	0.37(0.32,0.43)
		α_2	0.19(0.15,0.24)	0.21(0.17,0.27)	0.23(0.17,0.28)	0.23(0.17,0.30)	0.20(0.16,0.25)
		α_3	0.07(0.04,0.10)	0.08(0.05,0.12)	0.09(0.05,0.12)	0.09(0.05,0.14)	0.08(0.05,0.11)
vmy $\in [12,20]$	$SL \leq 80\,km/h$	α_0	0.41(0.35,0.47)	0.37(0.31,0.43)	0.38(0.32,0.44)	0.40(0.32,0.47)	0.39(0.34,0.45)
		α_1	0.43(0.37,0.49)	0.45(0.39,0.51)	0.43(0.37,0.49)	0.40(0.33,0.48)	0.43(0.37,0.49)
		α_2	0.10(0.07,0.14)	0.12(0.08,0.16)	0.12(0.08,0.17)	0.13(0.08,0.18)	0.11(0.08,0.15)
		α_3	0.06(0.03,0.09)	0.07(0.04,0.10)	0.07(0.04,0.11)	0.07(0.03,0.12)	0.06(0.04,0.09)
	$SL \geq 90\,km/h$	α_0	0.41(0.35,0.47)	0.37(0.30,0.43)	0.37(0.30,0.44)	0.39(0.30,0.46)	0.39(0.33,0.45)
		α_1	0.39(0.33,0.46)	0.40(0.34,0.47)	0.39(0.32,0.45)	0.36(0.29,0.44)	0.39(0.33,0.45)
		α_2	0.14(0.09,0.18)	0.15(0.11,0.21)	0.16(0.11,0.22)	0.17(0.11,0.23)	0.15(0.10,0.19)
		α_3	0.06(0.04,0.10)	0.08(0.04,0.12)	0.08(0.04,0.12)	0.08(0.04,0.14)	0.07(0.04,0.11)
		γ	1.00(1.00,1.00)	0.92(0.80,1.00)	0.86(0.50,1.00)	0.78(0.45,1.00)	0.94(0.88,0.99)

Table AIII
The estimated α_i, $i = 0,1,2,3$, and γ parameters calculated using the GIDAS *data set (mean, (2.5, and 97.5 percentiles in brackets)).*

age cohort		2012-2014	2015-2017	2018-2020	crash year group 2012-2020 (weighted)
18-44	S0	97.43(97.20,97.66)	97.11(96.85,97.36)	96.90(96.63,97.17)	97.13(96.88,97.38)
	S1	1.66(1.43,1.89)	1.87(1.62,2.13)	2.01(1.74,2.29)	1.86(1.61,2.11)
	S2	0.49(0.35,0.64)	0.55(0.39,0.73)	0.59(0.42,0.78)	0.55(0.39,0.72)
	S3	0.42(0.32,0.54)	0.47(0.35,0.60)	0.49(0.37,0.63)	0.47(0.35,0.59)
	S1+	2.57(2.34,2.80)	2.89(2.64,3.15)	3.10(2.83,3.37)	2.87(2.62,3.12)
	S2+	0.91(0.74,1.09)	1.02(0.83,1.22)	1.09(0.88,1.31)	1.01(0.83,1.22)
	P2	3.99(3.91,4.08)	4.50(4.45,4.55)	4.84(4.80,4.89)	4.47(4.44,4.51)
	P3	0.18(0.16,0.20)	0.19(0.18,0.20)	0.20(0.19,0.21)	0.19(0.18,0.20)
45-64	S0	97.56(97.31,97.81)	97.35(97.09,97.60)	97.14(96.87,97.41)	97.34(97.09,97.59)
	S1	1.61(1.37,1.86)	1.76(1.51,2.02)	1.89(1.62,2.17)	1.76(1.51,2.02)
	S2	0.45(0.31,0.61)	0.50(0.35,0.67)	0.53(0.37,0.71)	0.50(0.35,0.66)
	S3	0.38(0.27,0.49)	0.40(0.29,0.52)	0.43(0.31,0.56)	0.40(0.30,0.53)
	S1+	2.44(2.19,2.69)	2.65(2.40,2.91)	2.86(2.59,3.13)	2.66(2.41,2.91)
	S2+	0.83(0.66,1.01)	0.90(0.71,1.09)	0.96(0.77,1.17)	0.90(0.72,1.10)
	P2	3.90(3.73,4.06)	4.26(4.16,4.36)	4.59(4.50,4.68)	4.26(4.19,4.33)
	P3	0.14(0.11,0.17)	0.14(0.12,0.16)	0.15(0.14,0.17)	0.15(0.13,0.16)
65-74	S0	97.42(97.16,97.67)	97.05(96.78,97.31)	96.88(96.60,97.16)	97.09(96.83,97.35)
	S1	1.63(1.40,1.88)	1.87(1.61,2.14)	1.99(1.72,2.28)	1.84(1.59,2.11)
	S2	0.53(0.37,0.70)	0.60(0.42,0.79)	0.63(0.44,0.83)	0.59(0.42,0.78)
	S3	0.43(0.31,0.56)	0.48(0.35,0.63)	0.49(0.36,0.65)	0.47(0.35,0.62)
	S1+	2.58(2.33,2.84)	2.95(2.69,3.22)	3.12(2.84,3.40)	2.91(2.65,3.17)
	S2+	0.95(0.77,1.15)	1.08(0.87,1.30)	1.13(0.91,1.36)	1.06(0.86,1.28)
	P2	3.82(3.63,4.00)	4.38(4.26,4.49)	4.65(4.55,4.76)	4.31(4.24,4.39)
	P3	0.17(0.13,0.21)	0.19(0.16,0.21)	0.18(0.16,0.20)	0.18(0.17,0.20)
75+	S0	96.60(96.14,97.03)	96.27(95.89,96.64)	95.96(95.57,96.35)	96.29(95.95,96.64)
	S1	2.05(1.70,2.43)	2.24(1.90,2.59)	2.42(2.06,2.80)	2.24(1.91,2.58)
	S2	0.73(0.51,0.99)	0.80(0.56,1.07)	0.87(0.61,1.16)	0.80(0.56,1.06)
	S3	0.61(0.40,0.85)	0.69(0.48,0.92)	0.74(0.52,0.99)	0.67(0.47,0.89)
	S1+	3.40(2.97,3.86)	3.73(3.36,4.11)	4.04(3.65,4.43)	3.71(3.36,4.05)
	S2+	1.35(1.05,1.67)	1.49(1.20,1.80)	1.61(1.30,1.94)	1.47(1.19,1.77)
	P2	4.98(4.48,5.50)	5.42(5.13,5.72)	5.88(5.62,6.14)	5.42(5.22,5.63)
	P3	0.29(0.17,0.42)	0.34(0.26,0.42)	0.36(0.30,0.43)	0.32(0.27,0.37)

Table AIV
S-distribution for different time period and vehicle group [12,20] and age cohort. Note: P3 ≙ deceased.

Figures

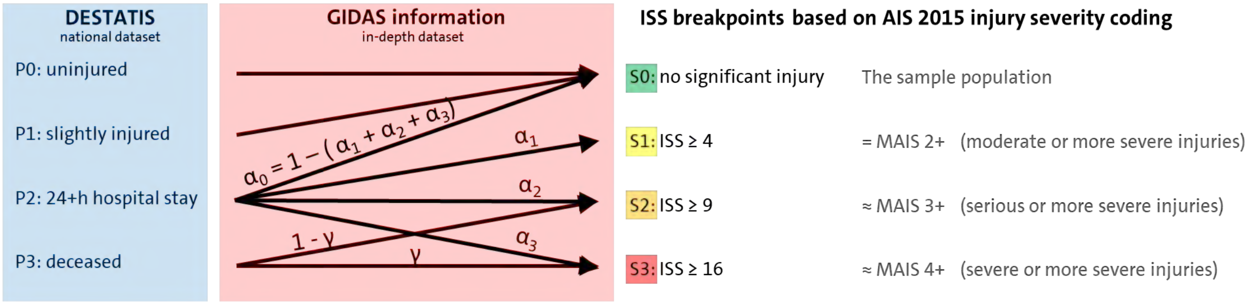

Fig. A2: Mapping of the P-scale to the S-scale.

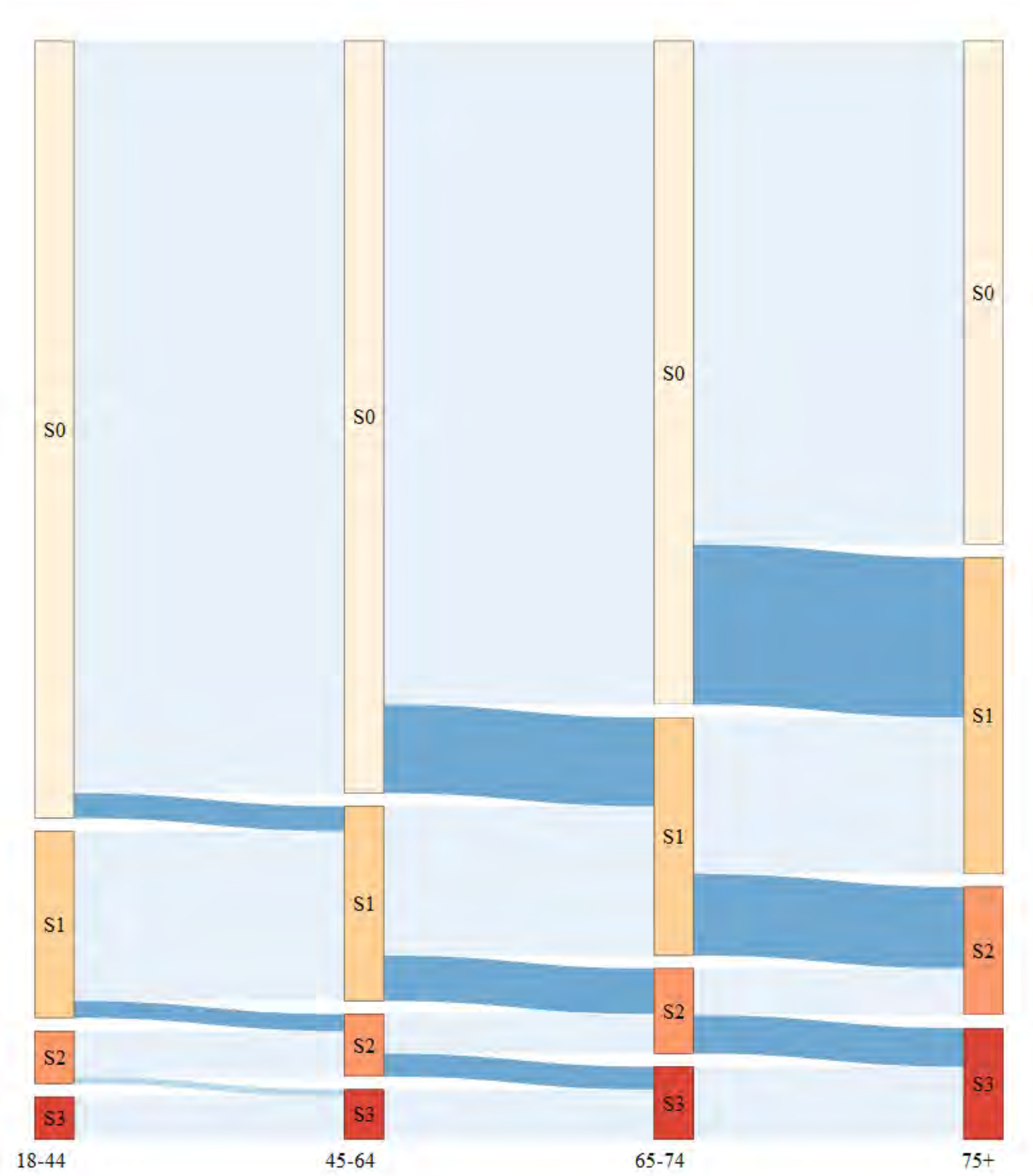

Fig. A3: Shift in injury severity ($S0 \rightarrow S1 \rightarrow S2 \rightarrow S3$) across the age groups (18–44, 45–64, 65–74, 75+ yoa), for vehicle model years (vmy) [12,20], and crash years 2012–2020 (weighted).
Flow: Dark blue: statistically significant shift (95 %-level), blue: non stat. sig. shift, light blue: no change in injury severity.
Columns represent the 10% of the crashes with the most severe injuries.

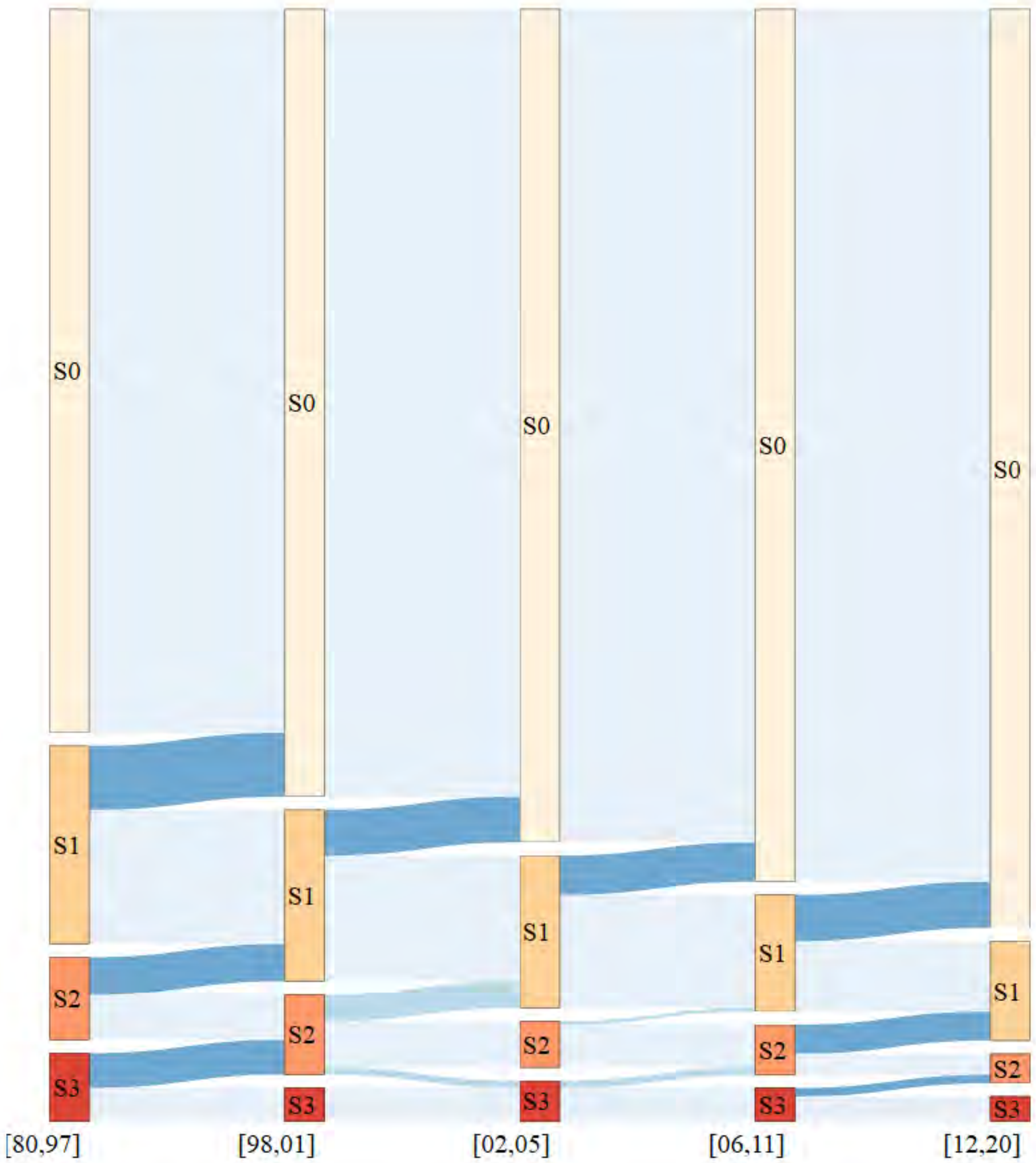

Fig. A4: Shift in injury severity ($S3 \rightarrow S2 \rightarrow S1 \rightarrow S0$) across vehicle model years (vmy) [80,97], [98,01], [02,05], [06,11], *to* [12,20], *for crash years 2012–2020 (weighted).*
All injury severity shifts are statistically significant at the 95%-level.
Columns represent the 10% of the crashes with the most severe injuries.

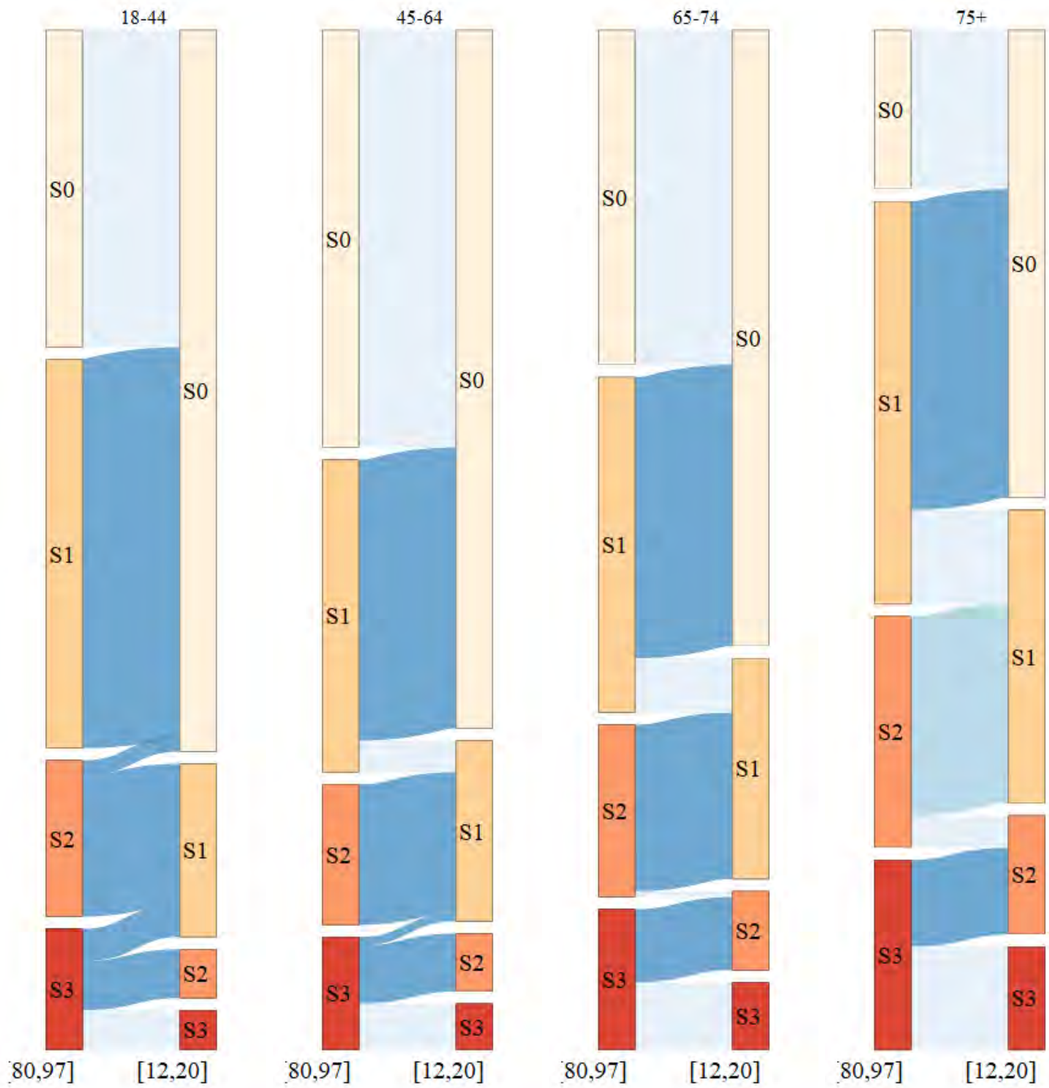

Fig. A5: Comparison of injury severity ($S3 \rightarrow S2 \rightarrow S1 \rightarrow S0$) between vehicle model years (vmy) $[80{,}97]$ to $[12{,}20]$, stratified by age group for (18–44, 45–64, 65–74, 75+ yoa), for crash years 2012–2020 (weighted). Flow: Dark blue: statistically significant shift (95 %-level), blue: non stat. sig. shift, light blue: no change in injury severity.
Columns represent the 10% of the crashes with the most severe injuries.

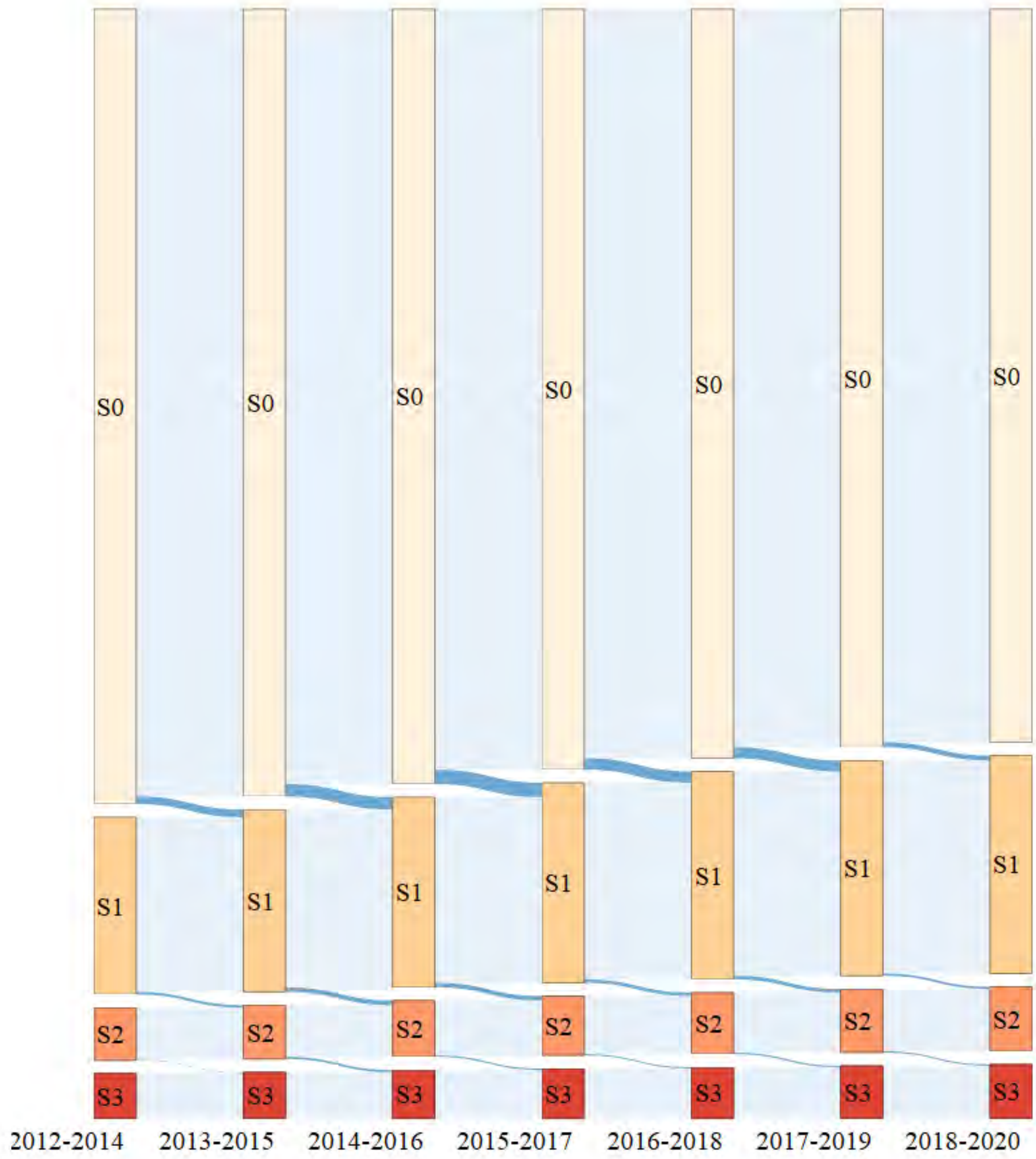

Fig. A6: Shift in injury severity ($S3 \rightarrow S2 \rightarrow S1 \rightarrow S0$) across the crash years (2012-2014, 2013-2015, 2014-2016, 2015-2017, 2016-2018, 2017-2019, 2018-2020), for vehicle model years (vmy) [12,20].
Flow: Dark blue: statistically significant shift (95 %-level), blue: non stat. sig. shift, light blue: no change in injury severity.
Columns represent the 10% of the crashes with the most severe injuries.

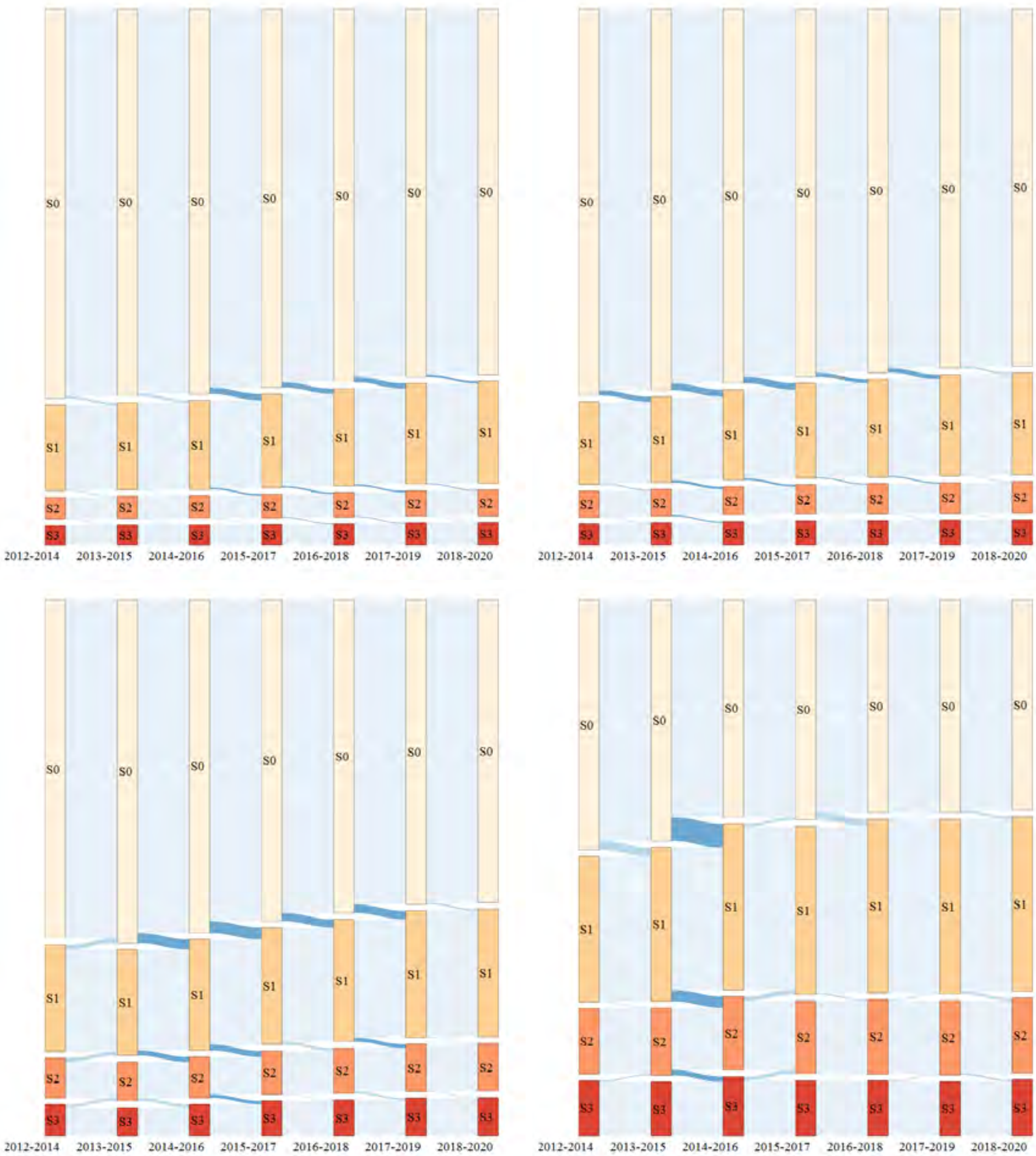

Fig. A7: Shift in injury severity ($S3 \rightarrow S2 \rightarrow S1 \rightarrow S0$) across the crash years (2012-2014, 2013-2015, 2014-2016, 2015-2017, 2016-2018, 2017-2019, 2018-2020), stratified by age group for (18–44, 45–64, 65–74, 75+ yoa), for vehicle model years (vmy) [12,20].
Flow: Dark blue: statistically significant shift (95 %-level), blue: non stat. sig. shift, light blue: no change in injury severity.
Columns represent the 10% of the crashes with the most severe injuries.

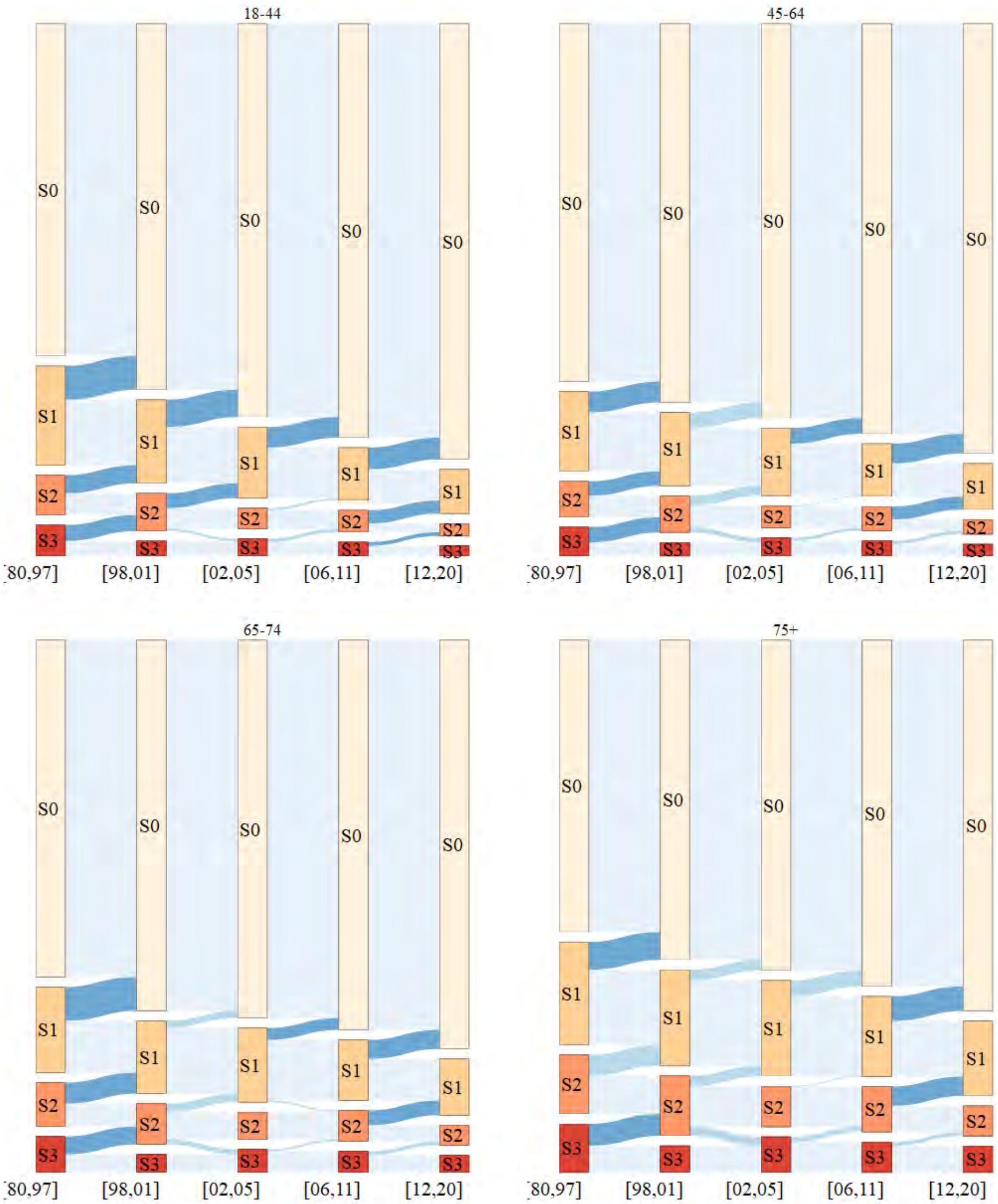

Fig. A8: Shift in injury severity ($S3 \rightarrow S2 \rightarrow S1 \rightarrow S0$) across vehicle model years (vmy) $[80{,}97]$, $[98{,}01]$, $[02{,}05]$, $[06{,}11]$, to $[12{,}20]$, stratified by age group for (18–44, 45–64, 65–74, 75+ yoa), for crash years 2012–2020 (weighted).
Flow: Dark blue: statistically significant shift (95 %-level), blue: non stat. sig. shift, light blue: no change in injury severity.
Columns represent the 10% of the crashes with the most severe injuries.

Temporal Stability of the α Parameter

The crash year is included as an independent variable in the logistic regression models of Section III (Vehicle Model Year) to test whether the α-parameters are temporally stable. The crash year is treated in a one-hot-encoding fashion where the crash year 2010 is the baseline to avoid any monotonic assumption. The following table states the p-values of the corresponding parameters in the three logistic regression models ($\alpha_i, i = 0,1,2$). For each α, the absolute maximum over all these coefficients can be used as a test statistic to test directly whether all 10 coefficients are different from zero. Bootstrap is employed for p-values to take the multiple testing issue into account. We obtain the following p-values $\alpha_0: 0.57, \alpha_1: 0.38, \alpha_2: 0.62$. Consequently, the logistic regression model seems unaffected by the crash year.

Crash Year	α_0	α_1	α_2
2011	0.22	0.43	0.38
2012	0.35	0.54	0.44
2013	0.30	0.23	0.13
2014	0.76	0.48	0.29
2015	0.17	0.35	0.17
2016	0.35	0.18	0.10
2017	0.32	0.12	0.67
2018	0.95	0.71	0.97
2019	0.54	0.15	0.87
2020	0.39	0.97	0.98

Table AV

p-values of the crash year coefficients in the logistic regression models for $\alpha_i i = 0,1,2$.

Injury Severity by Crash-Year and Age Group

Numerical Example

We simulated crash data, which mimics a national accident database and an in-depth accident study. The national accident database contained all the accidents but a lesser degree of detail for each crash. The in-depth accident study has a high level of detail for each accident but contains only a subset of all accidents and is biased due to some sampling criteria.

The General Approach

We consider four traffic domains which result in four different technical crash severity exposures. Additionally, we consider single and multi-vehicle crashes with one or more occupants in each vehicle. Injury risk functions transfer technical crash severity into medical crash injury severity. Medical crash injury severity is measured in two scales. One with a higher level of detail is denoted by IS and 5 levels are considered: 0,1,2,3,4. The levels are motivated by the MAIS: 0 (MAIS0); 1 (MAIS1); 2 (MAIS2); 3 (MAIS3); 4 (MAIS4+). The other scale is P and corresponds to a police-reported injury measure. It contains 3 levels $(0,1,2)$ which corresponds to 2 (fatal), 1 (hospitalised for crash related causes and non-fatal), 0 otherwise. Occupants can be belted or unbelted, corresponding to different underlying injury risk functions.

The Specifics

To elaborate, we consider n crashes. Let $TR_i, i = 1, \ldots, n$ be the technical crash severity for crash i. Technical crash severity is sampled independently from the exponential distribution, where the traffic domain results in different parameters for the exponential distribution. The four traffic domains are motivated by the speed limits (in km/h) $(0,30], (30,50], (50,80], (80,100]$. The parameters of the exponential distribution are given by fitting an exponential distribution to the Δv values (in km/h) in GIDAS for that traffic domain. The memorylessness property of the exponential distribution is used to take the sampling issue of GIDAS into account, and only Δv values greater than 20 km/h are considered for the estimation.

For the four traffic domains we compute:

$$\begin{aligned}(0,30]\colon TR_i &\sim \exp(0.11), i = 1, \ldots, 0.15n,\\ (30,50]\colon TR_i &\sim \exp(0.03), i = 0.15n + 1, \ldots, 0.7n,\\ (50,80]\colon TR_i &\sim \exp(0.06), i = 0.7n + 1, \ldots, 0.9n,\\ (80,100]\colon TR_i &\sim \exp(0.05), i = 0.9n + 1, \ldots, n.\end{aligned}$$

The four traffic domains have a share of 0.15,0.55,0.2,and 0.1, respectively. Let SVC_i indicate if accident i is a single vehicle crash and we set $SVC_i \overset{i.i.d}{\sim} Bernoulli(0.1), i = 1, \ldots, n,$ i.e., $\mathbb{P}(\text{single vehicle crash}) = 0.1$, see [13]. We set the expected number of vehicles involved in a multi-vehicle crash to 2.08, see [41]. Additionally, the number of involved vehicles is assumed to follow a Poisson distribution. Let $NVC_i = \text{number of vehicles in accident } i = 1 + (1 - SVC_i)(Z_i + 1), Z_i \overset{i.i.d}{\sim} Pois(0.08),\ i = 1, \ldots, n.$ We say that the number of occupants per vehicle follows a Poisson distribution, and the expected number of occupants per vehicle is set to 1.4 [42]. Let $OpC_i = \text{Occupants per accident } i = \sum_{j=1}^{NVC_i} 1 + O_{j,i}, O_{j,i} \overset{i.i.d}{\sim} Pois(0.4), j = 1, \ldots, NVC_i, i = 1, \ldots, n.$ We set seat belt usage to 97 % [43] and draw this using independent Bernoulli random variables. We say that all vehicles (and occupants) involved in a crash have the same technical crash severity. The medical injury severity is obtained independently for each occupant. For each occupant with technical risk x and belt status b, the medical injury severity in the IS scale is sampled according to $\mathbb{P}(IS = is|\text{technical risk} = x, \text{belt status} = b) = p_{is,b}(x), is = 0,1,2,3,4,$ where the probabilities $p_{is,b}(x)$ are obtained from the injury risk functions as follows. Let

$$R_{\beta_0,\beta_1}(x) = 1/\left(1 + exp\left(-\beta_0 + 10 - \beta_1 x - (10/1.3)\left(\log(\log(x+1)+1)\right)^{1/2}\right)\right).$$

$R_{\beta_1,\beta_2}(x)$ is a logistic risk function, and the term $(10/1.3)\left(\log(\log(x+1)+1)\right)^{1/2}$ is used to push the curve to zero near the origin. Then, for $is = 1, \ldots, 4, b = 0,1$

$$p_{is,b} = R_{\beta_{0,b,is},\beta_{1,b,is}}(x) \prod_{k=is+1}^{4} \left(1 - R_{\beta_{0,b,k},\beta_{1,b,k}}(x)\right),$$

$$p_{0,b}(x) = \prod_{k=1}^{4} \left(1 - R_{\beta_{0,b,k},\beta_{1,b,k}}(x)\right),$$

where the parameters $\beta_{0,b,is}, \beta_{1,b,is}, b = 0,1, is = 1,2,3,4$ are given by the following Table AVII.

g	1				0			
is	1	2	3	4	1	2	3	4
$\beta_{0,g,s}$	-3.52	-6.05	-8.45	-10.12	-1.57	-4.10	-6.32	-7.82
$\beta_{1,g,s}$	0.08	0.08	0.08	0.09	0.08	0.08	0.08	0.09

Table AVII: Model Parameters

The police-reported risk P is obtained from IS as follows:[2]

$$\mathbb{P}(P \geq 1 | IS = is) = \begin{cases} 0 & is = 0, \\ 0.05 & is = 1, \\ 0.65 & is = 2,, \\ 0.94 & is = 3, \\ 1 & is = 4 \end{cases}$$

and

$$\mathbb{P}(P = 2 | IS = is) = \begin{cases} 0 & is = 0, \\ 0 & is = 1, \\ 0 & is = 2,. \\ 0.04 & is = 3, \\ 0.63 & is = 4 \end{cases}$$

Lastly, to simulate the sample criteria of an in-depth accident study as GIDAS, we set $\mathbb{P}(\text{accident } i \text{ in in-depth sample} \mid \text{at least one occupant in accident } i \text{ with } IS \geq 1) = 0.05$ and $\mathbb{P}(\text{accident } i \text{ in in-depth sample} \mid \text{no occupant in accident } i \text{ with } IS \geq 1) = 0$. The small probability of 0.05 should consider that an in-depth study such as GIDAS is limited to only some areas of the national sample. Each accident in the in-depth accident study is then included in full detail, which means IS, P, technical risk, and traffic domain here. In contrast, the national accident database contains each accident but only limited information on the traffic domain and police-reported injury severity P for each occupant.

The objective is to estimate the injury distribution in detail: MAIS0–1, MAIS2, MAIS3, MAIS4+. We denote this as $S = 0, 1, 2,$ and 3, respectively. We compare the mapping approach present in this paper with the weighting approach. The mapping approach is applied as presented in Section III (The Link between P- and S-Scale). For the α_i parameters, the four traffic domains are grouped into two groups. The first three traffic domains form the first group, and the fourth one forms the second group. For the weighting approach, each case in the in-depth accident study gets a weight such that specific marginal distribution criteria are matched. This weighting can be done only to marginal distributions available in both samples. Thus, in this simulation, the distribution of the traffic domains between the weighted in-depth accident study and the national level are matched. We simulated $4\,000\,000$

[2] As mentioned above, the injury severity IS is motivated by MAIS. Hence, $\mathbb{P}(P = 2 | IS = s)$ can be understood as the fatality risk of MAIS0, MAIS1, MAIS2, MAIS3, and MAIS4+ injuries, respectively. Similarly, $(P \geq 1 | IS = is)$ can be understood as the risk of hospitalization or dying. The probabilities are oriented by the empirical probabilities in GIDAS.

crashes, and the in-depth study sampled of these about 72 000. The obtained injury severity distribution in the P-scale is given in Table AVIII. The sampling criteria cause the in-depth sample to be heavily biased and tend towards more severe crashes.

P	full-sample	in-depth sample
0	98.83	91.33
1	1.08	7.99
2	0.09	0.67

Table AVIII: Relative injury severity distribution in P-scale (simulation result)

The injury distribution of the full-sample, in-depth study, and estimation approaches are given in Table AV. The weighted in-depth sample is less biased but the covariate traffic domain is not strong enough to reduce the bias substantially. Additional information, such as technical crash severity, would be required to eliminate all bias. The projection by mapping approach can eliminate the bias for $\mathbb{P}(S=2)$ and $\mathbb{P}(S=3)$ because the corresponding conditional probabilities are unaffected by the sampling criteria. The bias for $\mathbb{P}(S=1)$ and $\mathbb{P}(S=0)$ is also substantially reduced although not completely since the probability $\mathbb{P}(S=0|P=0)$ is assumed to be 100%, but it is actually 99.67%. However, this conditional probability cannot be estimated without bias using the in-depth sample due to the sample criteria. Indeed, the conditional probability in the in-depth sample is 97.3%. What can be done is a sensitivity analysis on the assumption $\mathbb{P}(S=0|P=0)=1$. The sample criteria of the in-depth sample results in a bias towards more severe injuries and, consequently, an undersampling of $S=0$ injuries. Hence, the probabilities of the in-depth sample for $\mathbb{P}_{in-depth}(S=s|P=0), s=0,1,2,3$ can be used as a lower bound in a sensitivity analysis. That is we set $\mathbb{P}_{national}(S=s|P=0)=\mathbb{P}_{in-depth}(S=s|P=0)$ and redo the mapping. The results of this approach are denoted as mapping (SALW) in Table AV. This mapping (SALW) can be understood as a bound towards a more severe injury distribution, i.e., given the sampling criteria of the in-depth study, the injury distribution of the full sample cannot be more severe than mapping (SALW). Note that even mapping (SALW) is less biased than the weighted in-depth sample.

S	≈*MAIS*	full-sample	in-depth sample	weighting	mapping	mapping (SALW)
0	0–1	98.76	90.75	91.26	99.08	96.43
1	2	0.89	6.57	6.23	0.57	3.14
2	3	0.22	1.67	1.58	0.22	0.29
3	4+	0.14	1.01	0.94	0.14	0.14

Table AIX: Relative injury severity distribution in S-scale (simulation result)