The Costs of Traffic Accident Hotspots

Bernhard Gahr\textsuperscript{1}, Katherine Caves\textsuperscript{2}, Junhan Wen\textsuperscript{3}, Kevin Koch\textsuperscript{1}, Shu Liu\textsuperscript{2}, and Felix Wortmann\textsuperscript{1}

Abstract—Despite efforts to reduce them, traffic accidents continue to increase and bypass reduction targets. The costs of traffic accidents are enormous, killing 1.35 million people every year and costing 3% of most countries’ GDP. Recent research aims to target interventions at high-accident-density locations, called accident hotspots. New methods and technologies can systematically identify hotspots, but it remains unclear whether hotspots contribute to accident costs as well as volume. This paper investigates the monetary and human costs of accident hotspots. We analyze a dataset of all accidents from 2011 - 2017 in Switzerland. We identify hotspots, then analyze their contributions to traffic accident costs. We find that hotspot accidents are not different in monetary costliness or injury rates from non-hotspot accidents, so hotspots drive costs along with accident volume. However, hotspot accidents are less fatal, so hotspot targeting might not be best for fatalities. If hotspots are reduced to normal road conditions, total monetary costs can be reduced by up to 5% per year as a theoretical upper bound. Targeting the top 10% most frequent, costly, injurious, or deadly hotspots yields different results for different cost types, with accident number and monetary cost targets creating the highest reductions overall.

I. INTRODUCTION

Despite continuous investment in road and vehicle safety, as well as improvements in technology standards, the number of road traffic accidents has steadily increased in recent decades. For example, the USA saw a 5.6% increase in the number of traffic-accident deaths from 2015 to 2016, on top of an 8.4% jump from 2014 to 2015. Road deaths in the European zone also increased in both 2015 and 2016, widening the gap to that region’s stated goal of halving road fatalities from 2010 to 2020 [1].

According to the World Health Organization, traffic accidents have heavy costs: 1.35 million people die every year from road accidents, and they cost most countries 3% of GDP [2]. Countries and international bodies go to great lengths to prevent and reduce the costs of road accidents [3], with a major focus on identifying and intervening upon the most vulnerable and high-risk people, places, and situations [2].

Certain locations are quantitatively more dangerous than others, for reasons ranging from poor road infrastructure to bad lighting or vulnerability to weather conditions. These high-risk locations are known as “accident hotspots” [4].

Accident hotspots are a natural point of intervention for reducing the number, severity, and costliness of traffic accidents. A key issue for this purpose is identifying hotspots systematically, then leveraging that information to reduce accidents. For example, warning drivers of upcoming hotspots can reduce the number of critical driving maneuvers like sharp braking [5].

However, although hotspots are a clear point of intervention, it is not clear how much they contribute to traffic accident costs. They probably generate significant costs due to sheer volume, but they might be outweighed by extremely high-cost accidents in non-hotspot locations. This paper investigates whether accident hotspots are major drivers of total traffic accident costs. We examine monetary costs as well as human costs, represented by injuries and fatalities. We also assess the potential for cost reduction from hotspot-based intervention strategies. This holistic view of hotspots’ impact can help direct how and where efforts should be spent to efficiently improve road safety.

II. RELATED WORK

This paper brings together two strands of literature. The first develops methods for identifying traffic accident hotspots, and the second examines the costs of traffic accidents. In this section, we briefly review those two strands and the small group of other studies that examine the costs of hotspot-like locations, times, seasons, and behaviors.

A. Hotspot Identification and Analysis

Over the past sixty years, various researchers have investigated road accident hotspots and developed a number of hotspot identification (HSID) methods.

Simple and common non-spatial techniques use roads’ underlying structure and consider traffic accidents on set road segments. The Crash Frequency (CF) method is simplest, determining segment danger from accident numbers while taking road conditions into account [6]. Like CF, the Crash Rate (CR) method uses accident numbers, but unlike CF, CR accounts for traffic volume [7]. Although both methods are intuitive and commonly used, they both fail to consider random fluctuations in accident numbers and are weaker statistically than other HSID methods [8].

Newer methods use more robust statistical analyses to identify accident hotspots. One of the most well known and widely used is the Empirical Bayesian (EB) method. EB statistically outperforms other HSID techniques, including CR and CF, but like those methods it fails to account for spatial data and accident patterns [9], [10].
Spatial data analysis methods are possible thanks to developments in Geographic Information Systems (GIS), precise geo-coded data, and digital maps [11], [10]. The most common spatial HSID techniques are K-means clustering, spatial autocorrelation, and Kernel Density Estimation (KDE). KDE is the strongest in terms of statistical performance — on par with EB — and has been researched extensively [12]. KDE is particularly useful for HSID when combined with other methods like repeatability analysis [13], statistical analysis [14], and K-means clustering [15].

The most recent method is a data-mining clustering technique known as Density-Based Spatial Clustering of Applications with Noise (DBSCAN). This classifies elements into clusters such that each cluster has higher element density than the area around it. DBSCAN can efficiently identify clusters of random shapes and discriminate between cluster members and outliers [5], [4].

B. Costs of Traffic Accidents

There is an extensive literature on estimating the costs of traffic accidents in specific countries [16], [17], [18], under specific conditions like seasonal weather [19], and in conjunction with other public health issues like alcohol use [20]. Researchers address total costs, the proportion of costs made up from specific accident types [21], and the difference between traffic accident costs and the potential costs of other scenarios like increased walking or cycling [22].

The costs of traffic accidents can be direct and indirect. Direct costs include property damage, medical care, emergency response, and insurance administration. Indirect costs mainly include lost productivity due to injuries, fatalities, and time lost in traffic jams [23]. Both cost types affect involved parties, third parties, and insurance. High-income countries are especially affected by indirect costs due to their higher potential losses in productive capacity and quality of life [2].

Calculating the cost of an incident like a traffic accident is complex. Direct costs seem straightforward, but medical costs are ongoing and emergency response costs like police and medical transportation are spread over many events. Indirect costs are even more challenging, requiring estimations of lost productivity with great uncertainty (i.e. a fatality in a young person), over large and diverse groups (i.e. the value of one hour spent in traffic for 50 people), and with unclear parameters (i.e. an injured person who has returned to work but is less productive than before). Methods for approximating costs vary, but most take an economic or epidemiological approach that includes some elements of cost-benefit analysis [22], cost-of-illness studies [23], [24], or willingness-to-pay studies [25].

C. The Cost of Hotspots

Reducing accidents by reducing high-risk situations is an obvious step, and many of the papers cited above make at least some recommendations that apply to reducing the biggest risk factors for traffic accidents in order to reduce their impact. However, new HSID techniques have created a new way of finding high-risk traffic situations and those have not been fully evaluated in terms of cost. The question remains whether hotspots found by HSID are major contributors to total traffic costs, or if they mainly drive accident numbers.

[26] addressed the issue, using EB and regression methods to identify “blackspots” and finding that well-applied safety policies can reduce up to 5% of total accident costs. [27] showed that hotspots predicted by a multivariate Poisson-lognormal (MVPLN) model correlate with the severity of observed accidents. [28] attempted to predict severe accidents specifically, finding that different types of machine learning techniques predict accidents of different severity levels. This indicates that different HSID methods might yeild hotspots with different impacts on costs.

Most of the research on hotspots focuses on prediction accuracy rather than cost impact. Conversely, research on the cost of accidents looks for the highest-cost accidents or accident types instead of hotspots. When cost studies look at high-frequency zones, they are identified very differently from how hotspots are found. This study combines those two approaches to examine the cost of hotspots and their impact on total costs.

III. DATA & METHODOLOGY

We use the DBSCAN method to identify accident hotspots, following the DBSCAN application and Swiss coordinate system transformation procedure used on the same dataset in [4]. The DBSCAN algorithm takes two parameters for clustering: the minimum number of points within one cluster and the minimum distance between two points in the same cluster. Following [4], we set the parameters of the DBSCAN algorithm to a minimum number of 10 accidents per hotspot in moving five-year periods, or a minimum of two accidents per year, and a minimum distance of 15 meters between two accidents. We did not apply different HSID mechanisms in this case because most of the more advanced methods require controlling for traffic density, which was not available in the data used here.

A. Data

The major challenge of assessing hotspots’ contribution to traffic accident costs is finding a dataset rich enough to identify hotspots, calculate direct costs, and approximate indirect costs. We use a dataset of all Swiss car accidents 2011-2017, which includes all of the necessary spatial data for HSID, on-scene estimates of total accident cost from law enforcement officers, and injury and fatality counts for each accident.

The estimate of financial costs comes from on-scene police officers, who give the “total predicted cost” of the accident in their incident reports. We refer to these as monetary costs. Human costs include the number of injured people and fatalities in each accident. We never combine injuries and fatalities because we do not have sufficient data to determine the cost ratio between injury and death.

Data were collected by on-scene Swiss policemen at the time of each accident using their standard accident reporting
protocol. In addition to the spatial data for DBSCAN HSID, each accident report includes the date, total predicted cost (damage, in thousands of Swiss Francs or kCHF), number of injuries, and number of fatalities. Table 1 shows descriptive statistics. The full sample includes 377,648 total accidents, accounting for more than 2.9 billion CHF in monetary damages and 163,972 injuries plus 1,870 fatalities.

### IV. Results

We begin by describing the accident hotspots we found and their contributions to accident numbers. Then we test whether hotspots are costlier, less costly, or the same cost as non-hotspot accidents. If hotspot accidents are similar to or more costly than non-hotspot accidents, then hotspots are a good place to intervene for the goal of reducing traffic accident costs. We finish by identifying the potential cost savings from intervening on hotspots.

#### A. Hotspots

We identify between 1,260 and 1,400 accident hotspots per five-year period (Figure 1, left axis), accounting for 8-9% of all traffic accidents (right axis). Over time, the number of accident hotspots increases, as does the proportion of accidents they account for. This indicates that hotspots are an important point of intervention for reducing traffic accident numbers. However, hotspots vary in their individual impact.

**Figure 2** shows the distribution of hotspots by size in terms of accident numbers for samples ending in 2016, 2017, and 2018. The distribution is heavily right-skewed, with most hotspots being smaller and fewer having large accident numbers. For the 2018 cohort, the mean number of accidents per hotspot is 17.3. The median hotspot size is 13 accidents, with only the top four hotspots having more than 100 accidents per five-year time period.

In every time period, there is an outlier hotspot with more than 400 accidents. Over time, the size of hotspots increases — the time period ending in 2018 has hotspots in the 140- and 160-accident ranges while the other time periods have gaps between 100 and 400 accidents per hotspot.

#### B. The Monetary Cost of Hotspots

First, we test whether the monetary costs of hotspot accidents are different from non-hotspot or overall average accidents. Figure 3 shows the mean cost of accidents overall, in hotspots, and outside hotspots over time. The shaded areas around the lines are one standard deviation from the mean. We find that hotspot accidents are not significantly different from the mean or from non-hotspot accidents at any significance level: they make up for their “fair share” of monetary costs. The trend over time appears to be roughly parallel, with hotspot accidents following the cost trends of non-hotspot and average accidents. This confirms that intervening on a hotspot should, on average, reduce costs proportionate to the number of accidents eliminated.
C. The Human Cost of Hotspots

Second, we test whether the human costs of hotspot accidents differ from non-hotspot or average accidents. Figure 4 shows the human cost of hotspot accidents by comparing the proportion of injuries and fatalities that occur inside hotspots to the proportion of accidents that are in hotspots, with one standard deviation from the mean shaded. We do not show a 95% confidence interval to avoid muddying the figure, but fatalities are significantly different from the rate of accidents at hotspots at the 0.05% significance level. If the proportion of injuries or fatalities in hotspots is significantly different from the incidence of hotspot accidents, then hotspots make up for more or less than their “fair share” of human costs. If not, hotspots are similar in human costs to non-hotspot accidents.

We find that hotspot injury rates are not significantly different from the rate of hotspot accidents, so hotspots account for their fair share of injuries. However, the proportion of total fatalities that occurs in hotspots is significantly lower than the proportion of accidents that are hotspots, so hotspots have lower fatalities than non-hotspot accidents. This indicates that hotspots are important for the bulk of human costs, but not as important for the most extreme human costs.

Fig. 4. Human cost contribution by accidents at hotspots, injured people at hotspots and fatalities at hotspots including the 1st standard deviation interval.

D. Potential Savings

As previously discussed, we need to be cautious when deriving implications from our results because the data does not perfectly capture costs as a formal cost-of-illness study would. What we can do is establish bounds within which potential cost savings from interventions at hotspots might lie. To do this, we explore what would happen to costs if hotspots were reduced to the point that they no longer qualified as hotspots.

There were 1,384 hotspots in the cohort ending in 2018. These hotspots account for 8.9% of all accidents, or 23,960 total accidents. The average monetary cost per accident in this group is 8,204 CHF, making a total cost of 40,970,513 CHF for hotspot accidents in the most recent time period, out of the total 430 million CHF.

Eliminating a hotspot completely is the best-case scenario for traffic intervention, but that does not mean there are zero accidents at that location. Because of the lower limit for hotspot status, elimination only reduces accidents at each location to two per year for five years. If we reduce every hotspot to two accidents, 2,768 accidents remain for a total hotspot-accident monetary cost of 18,261,841 CHF per year. Using the total annual monetary costs of traffic accidents in 2018 as a benchmark, this is a 5.27% savings and an absolute upper bound. Interestingly, 5% is the same number that [26] reached in their analysis.

Using similar calculations on injury and fatality data, as shown in Equation (1), we can find the potential human cost savings. In Equation (1), the first term returns the number of injuries that happened at a hotspot. The second term shows the potential savings with the lower limit of two accidents per year remaining. Applying this formula the hotspot data, we find potential human cost savings of 860 injuries and 2.2 fatalities per year.

\[
\text{Reduction} = \left( \frac{\text{Injuries}_{\text{total}} - \text{Injuries}_{\text{at HS}}}{\text{Accidents}_{\text{at HS}}} \right)
\]

To highlight a more realistic scenario, we can use the same hotspot elimination calculations to find the upper bound if an intervention eliminated the top 10% of hotspots by frequency, monetary cost, injuries, and fatalities. Table 2 shows the annual reduction for each cost type if those hotspots are reduced to normal accident levels of two accidents per year. All numbers are the average of the most recent five-year cohort.

<table>
<thead>
<tr>
<th>Priority</th>
<th>Savings</th>
<th>Accidents</th>
<th>£CHF</th>
<th>Injuries</th>
<th>Fatalities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accidents</td>
<td>1,013</td>
<td>7,202</td>
<td>200</td>
<td>-0.04</td>
<td></td>
</tr>
<tr>
<td>(%)</td>
<td>1.88%</td>
<td>1.74%</td>
<td>0.84%</td>
<td>(-0.02%)</td>
<td></td>
</tr>
<tr>
<td>Monetary</td>
<td>729</td>
<td>11,053</td>
<td>205</td>
<td>-0.24</td>
<td></td>
</tr>
<tr>
<td>(%)</td>
<td>1.35%</td>
<td>2.66%</td>
<td>0.86%</td>
<td>(-0.10%)</td>
<td></td>
</tr>
<tr>
<td>Injuries</td>
<td>550</td>
<td>6,239</td>
<td>390</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>(%)</td>
<td>1.02%</td>
<td>1.50%</td>
<td>1.65%</td>
<td>(0.15%)</td>
<td></td>
</tr>
<tr>
<td>Fatalities</td>
<td>382</td>
<td>2,970</td>
<td>101</td>
<td>4.16</td>
<td></td>
</tr>
<tr>
<td>(%)</td>
<td>0.71%</td>
<td>0.72%</td>
<td>0.43%</td>
<td>(1.72%)</td>
<td></td>
</tr>
</tbody>
</table>

By adjusting the top 10% of hotspots to normal road safety levels, we can reduce cost elements by up to 2.66%. In general, targeting the highest accident-number and highest monetary-cost hotspots creates the biggest percent reductions. Reductions range from 1.65% to 2.66% in the targeted cost factor, and are generally positive between .43% and 1.74% in non-target cost factors except for fatalities, which are low and even slightly negative in two cases. Overall, accident numbers, injuries, and monetary costs reduce in response to any target and the most when targeted themselves.
The difference with fatalities may be that they are somehow unique relative to the other cost factors, or it may arise from the relatively low number of fatalities and the influence of a few very high-fatality accidents. We cannot empirically prioritize or directly compare across categories because we do not have the information necessary to compare a fatality to a cost savings in another category. Decision-makers can identify and enact their own priorities.

V. Discussion and Outlook

The newest HSID methods can systematically identify traffic accident hotspots. These are an attractive target for intervention because they can offer efficient impact on accident frequency. The main goal of intervening at traffic hotspots is to reduce the total cost of traffic accidents, in terms of both monetary and human losses. This study has shown that hotspots are a good targeting strategy for reducing monetary costs and injury rates associated with traffic accidents, although they may not be the best approach for reducing fatalities.

A. Limitations

This study has a number of limitations. Our measures for costs are not perfectly aligned with the direct and indirect costs usually used to assess traffic accidents. In addition, our monetary costs are estimates from on-scene law enforcement officers, not final accounting. We use injury rates and types to approximate lost productivity and medical costs, but these are neither precise nor necessarily comparable across categories. We intentionally keep injuries and fatalities separate, and never mix either with monetary costs, because we do not have the data to adequately value injuries or fatalities in monetary terms.

Our dataset does not include any information for third-party costs like loss of productivity in traffic or insurance administration. There may be some indication of those costs in police officer’s reporting because they can observe traffic and vehicle damage, but none of our estimates can say anything about costs except approximate monetary costs incurred during the accident, injury rates, and fatality rates.

The most important limitations are on our estimates of the theoretical upper bound of cost savings if hotspots are reduced to normal accident levels. We cannot emphasize enough that these are the absolute highest possible savings and that they are theoretical. There may be an unavoidable reason for a hotspot that mean it cannot ever be brought to normal accident levels. Therefore, these estimates should never be an expectation or even a goal. They are merely the maximum potential.

B. Contributions

This study makes three groups of contributions. First, we identify hotspots in Switzerland and show that the number and size of accident hotspots is increasing over time. Second, we demonstrate that hotspots are good intervention targets by showing that their costs and injury rates per accident are similar to non-hotspot accidents. We also show that hotspots are not the best method of targeting traffic fatalities. Finally, we test the potential of using hotspots to reduce traffic accident costs through total and targeted hotspot remediation.

When we look at the trends in accidents, hotspots, and hotspot sizes over time, we find that all three are increasing in Switzerland. As accident numbers grow and continue to consolidate into hotspots, the hotspots themselves may become even more important for managing increasing traffic problems, accidents, and costs.

The monetary costs of accidents in hotspots are not different from accidents outside hotspots. Therefore fixing hotspots to reduce accident frequency will lower costs proportionally — on average. That finding can be applied in other countries that may not have the right data available to support the assumption that reducing hotspots will reduce total traffic accident costs. However, the data in this study comes from Switzerland, where non-hotspot accidents may be particularly common and costly due to the country’s many mountainous areas.

The human costs of accidents are more nuanced. Rates of injury in hotspots are not different from the rate of total accidents in hotspots, so reducing hotspots should reduce injuries. Conversely, rates of fatality in hotspots are lower than expected, so hotspot accidents are not as deadly as non-hotspot accidents. Injuries are much more common than fatalities and make up for the vast majority of the human-cost-related incidents, but fatalities are the worst possible outcome so they can never be overlooked.

If we test hotspot elimination as a route to reducing traffic-accident costs, we find promising potential. Reducing all hotspots to non-hotspot levels (two accidents per year every year for five years) can save as much as 5% of total monetary costs per year, 3% fewer injuries, and 1% fewer fatalities per year. If policymakers spot-reduce the top 10% of hotspots based on the hotspots’ accident numbers, monetary cost, injury rate, or fatality rate, they can reduce those up to 1.88%, 2.66%, 1.65%, or 1.72%, respectively. However, choosing a priority means that cost factor is reduced more than the others, and fatalities require the most specific targeting. This research can explore the potential effects of targeted policies, but cannot recommend the best targets.

C. Outlook

As HSID improves, traffic accident hotspots have the potential to become a useful tool for reducing traffic accident costs. The information that hotspots are important cost drivers justifies their use as a tool for planning accident-prevention actions. Further research can look into third-party costs, determine whether the trends found here are consistent with other places and times, and deepen the analysis of injury levels, fatality prevention, and hotspot selection for optimal intervention. Some hotspot-based interventions are already underway at road authorities in Switzerland, and future research should investigate the impact of these approaches on all types of traffic accident costs.