



## **Exploring turn signal usage patterns in lane changes: A Bayesian hierarchical modelling analysis of realistic driving data**

Downloaded from: <https://research.chalmers.se>, 2025-12-10 01:15 UTC

Citation for the original published paper (version of record):

Jokhio, S., Olleja, P., Bärghman, J. et al (2024). Exploring turn signal usage patterns in lane changes: A Bayesian hierarchical modelling analysis of realistic driving data. IET Intelligent Transport Systems, 18(2): 393-408.  
<http://dx.doi.org/10.1049/itr2.12457>

N.B. When citing this work, cite the original published paper.

# *IET Intelligent Transport Systems*

## Special issue Call for Papers

---

**Be Seen. Be Cited.  
Submit your work to a new  
IET special issue**

Connect with researchers and experts in your field and share knowledge.

Be part of the latest research trends, faster.

**Read more**



The Institution of  
Engineering and Technology

# Exploring turn signal usage patterns in lane changes: A Bayesian hierarchical modelling analysis of realistic driving data

Sarang Jokhio<sup>1</sup>  | Pierluigi Olleja<sup>2</sup> | Jonas Bärgrman<sup>2</sup> | Fei Yan<sup>1</sup> | Martin Baumann<sup>1</sup>

<sup>1</sup>Department of Human Factors, Institute of Psychology and Education, Ulm University, Ulm, Germany

<sup>2</sup>Division of Vehicle Safety, Department of Mechanics and Maritime Sciences, Chalmers University of Technology, Gothenburg, Sweden

## Correspondence

Sarang Jokhio, Department of Human Factors, Institute of Psychology and Education, Ulm University, 89081 Ulm, Germany.  
Email: sarang.jokhio@uni-ulm.de

## Funding information

H2020 Marie Skłodowska-Curie Actions, Grant/Award Number: 860410

## Abstract

Using turn signals to convey a driver's intention to change lanes provides a direct and unambiguous way of communicating with nearby drivers. Nonetheless, past research has indicated that drivers may not always use their turn signals before starting a lane change. In this study, realistic driving data are analyzed to investigate turn signal usage during lane changes on highways in and around Gothenburg, Sweden. Turn signal usage is examined and factors that influence it are identified by employing Bayesian hierarchical modelling. The study found that drivers used their turn signal before changing lanes in 60% of cases, after starting the lane change in 33% of cases, and did not use it at all in 7% of cases. The Bayesian hierarchical modelling results indicate that various factors, such as the speed and direction of lane changes and the presence of surrounding vehicles, influence the usage of turn signals. The study concludes that understanding the factors affecting turn signal usage is crucial for improving traffic safety in current and future mixed traffic with autonomous vehicles. The study discusses the implications of findings concerning increasing turn signal compliance through general policy-making, improving existing in-vehicle technologies and including turn signal usage in Pay-As-You-Drive insurances.

## 1 | INTRODUCTION

The behaviour of drivers in traffic can be unpredictable, leading to complex and dynamic interactions that pose a challenge to the smooth functioning of autonomous vehicles (AVs). To be able to navigate and respond to human drivers safely and efficiently, AVs should be programmed with the ability to anticipate and interpret human driver behaviour. Understanding and interpreting the various explicit and implicit signals that human drivers use in traffic is crucial for AVs to communicate effectively and interact with other road users in different situations. One of the interaction situations that can occur is during a lane change. The lane change is a frequently performed but complex manoeuvre that requires effective communication with other drivers to ensure safe completion. In a lane-changing scenario, a driver uses implicit and explicit forms of communication to convey the intentions to other road users [1]. An implicit or indirect form of communication during a lane change occurs when the driver positions the vehicle in a certain way, such as moving closer to the lane boundary line, as an indication of their

intention to change lanes. An explicit or direct form of communication during a lane change is using a turn signal to announce the driver's intentions to change lanes to surrounding drivers.

A turn signal is a crucial safety feature that plays a vital role in effective communication between drivers, as it allows them to announce their intention to change lanes [2]. As such, it is essential that turn signals are used properly and consistently, particularly before starting a lane change. Although the literature lacks a formal definition of "proper turn signal use" it is widely recognized that it involves activating the turn signal before starting a lane change. Using a turn signal before changing lanes allows drivers in the target lane to anticipate and prepare for the upcoming lane change by providing them with enough time to react and adjust their speed or position accordingly. Many countries require drivers to use turn signals before making a lane change, but there is a lack of clear regulations regarding the onset time (i.e. when it should be activated). For example German road traffic law only requires drivers to signal the lane change intention "in good time" [3], leaving room for interpretation. However, some states in the United States,

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *IET Intelligent Transport Systems* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology.

such as California, have more specific regulations in road traffic law, requiring drivers to use a turn signal at least 5 s before a lane change [4]. Proper turn signal usage is essential, as it could avoid many potential crashes [2] and also increase cooperative behaviour (such as opening up gaps for lane changing vehicles) by drivers in the target lane [1, 5]. The use of proper turn signals is also imperative in the context of future mixed traffic scenarios, where AVs will coexist with conventional human-driven vehicles on shared roadways. This is because, as AVs rely heavily on strict rules, they may not understand the implicit forms of communication used by human drivers [6].

## 1.1 | Relevant literature

Research on lane changes has substantially increased in the past two decades due to advancements in data collection technology and the availability of large-scale data [7, 8]. However, a substantial portion of lane change research focuses on characteristics like lane change duration [7, 9, 10], gap acceptance [11–13]. Despite its crucial role in safe lane-changing, turn signal usage has remained relatively under-explored. We have summarized the findings of studies investigating turn signal usage in the following section.

Tijjerina et al. [14], investigated turn signal usage of drivers while performing a lane change. The data for the study was collected on public roads in Ohio, USA, using an instrumented vehicle, with the driver being accompanied by an observer. They reported that a turn signal was used 92% of the time. However, the observers' presence in the passenger seat might have resulted in higher compliance than if no observers were present [15]. In a later study by Lee et al. [15], 16 participants drove an instrumented vehicle on an interstate highway in Virginia, USA, without an observer. The study found that the turn signal was used only 44% of the time. Turn signal usage for lane changes to the right was only 35% compared to 48% in the case of lane changes to the left lane. Unlike the above two studies, Ponziani [2] used an observer with two tally counters in a vehicle driving around Ohio, USA. The observer recorded the lawful use of the turn signal on one counter and unlawful or neglected on the second tally counter. A total of 2000 data points were gathered for lane changes. It was found that only 51.65% of drivers complied with proper turn signal usage while making a lane change. Lin and Bao [16] used data collected in Michigan, USA, during the Safety Pilot field-operational-test sponsored by the National Highway Traffic Safety Administration [17]. They reported the use of turn signals in about 70% of lane changes. Fitch et al. [18] conducted an analysis of lane-change events using data from the 100-Car Naturalistic Driving Study. Of the total 135 lane change cases, three were crashes, and the remaining were near-crashes. The study found that while 85% of drivers used their turn signals during planned left-lane changes, only 24% did so when making unplanned left-lane changes to avoid a forward crash threat.

All of the above studies were from the USA. However, similar numbers were reported in the two studies conducted in China. For example, Dang et al. [19], used an instrumented vehicle and

studied 12 participants who drove approximately 200 km (on each side) on a two-lane highway between the city of Beijing and Baoding. On the first leg of the journey (Beijing to Baoding), the drivers drove the vehicle with their usual driving behaviour. On the second leg of the journey (Baoding to Beijing), drivers were asked to make frequent lane changes and use turn signals before starting a lane change. The percentage of turn signal usage in the first leg was about 40%. However, despite being instructed to use the turn signal before starting a lane change, the study found that its usage in the second leg of the drive was only about 65%. The second study by Wang et al. [20], used data collected during the Shanghai Naturalistic Driving Study. They analyzed lane changes made by the vehicle in the adjacent lane (both left and right) to the instrumented vehicle. After extracting the time of lane changes (called cut-in the original paper), they observed turn signal usage by watching forward roadway videos from the subject vehicles. The results showed that turn signal usage was less than 50% on all types of roads, including freeways and expressways. These results indicate that turn usage during lane change is relatively low, even though traffic law mandates it in many countries worldwide.

## 1.2 | Current study

As previously emphasized, the use of turn signals to indicate intent during a lane change is a vital aspect, yet it has received less attention than other characteristics, such as the duration of the lane change and acceptance of the gap. Only a few studies on this topic have been conducted, and they predominantly originate from just two countries: the United States and China. To the authors' knowledge, there is a lack of research on turn signal behaviour, specifically in the context of drivers from European countries. The current literature also primarily relies on observations (turn signal used by other vehicles), by a person in the vehicle or via videos recorded by instrumented vehicles. Such methodologies are prone to errors, particularly when events are observed or recorded from the perspective of the ego vehicle, rather than focusing on the behaviours of the ego vehicle itself. For example there is a scarcity of research that differentiates between turn signals used before and after starting a lane change, except for [16]. This scarcity might largely be due to the challenges posed by observation-based methodologies, which make it difficult to accurately determine the relative timings of turn signal activation in comparison to the actual initiation of the lane change. Additionally, there is a gap in extensive research exploring the various factors that could potentially impact the usage of turn signals during lane changes. Therefore, this study aims to examine the usage of turn signals among Swedish (European) drivers by analyzing data on their naturalistic behaviour during lane changes. Additionally, the paper provides an in-depth analysis of potential influencing factors on turn signal usage.

We used Bayesian hierarchical modelling to analyze the turn signal usage data. We used realistic driving data, which typically has a hierarchical or multilevel structure where observations are nested within participants. In recent years, Bayesian approaches

have been increasingly applied in transportation and other fields such as social and behavioural sciences [21–23]. In the transportation domain, the Bayesian approach has been used to study road safety [24], such as crash estimation, prediction of road accidents, and road network safety evaluation [25–28]. In addition, Bayesian techniques have been employed to develop reference models [29] and to produce synthetic data that mimics driver behaviour [30].

The structure of this paper is as follows. In Section 2, we thoroughly examine the data utilized in this study. In Section 3, we discuss Bayesian hierarchical modelling, followed by a presentation of our findings in Section 4. Lastly, the implications of our results, potential limitations, and recommendations for future research are discussed in Section 5.

## 2 | DATA DESCRIPTION

The data used in this paper were collected during the L3Pilot Project [31]. A dataset collected by Volvo Car Corporation for this project includes vehicles equipped with an automated driving function being driven on the highway surrounding the city of Gothenburg, Sweden. This dataset includes both trips where the driving task was purely manual (baseline) and trips where the automated driving function was activated (treatment). Drivers in the baseline dataset were primarily non-professional drivers. The professional drivers were involved in part of the baseline and in the whole treatment dataset. In the context of L3Pilot, the term “professional” referred to drivers with a qualification in operating prototype vehicles for testing specific features [31, 32]. Professional drivers were also compensated for driving. Non-professional drivers, on the other hand, are individuals who are recruited within the company [31]. In our study, only baseline trips with both non-professional and professional drivers were considered. The cars were equipped with cameras facing outside towards traffic (Front and Rear) and inside towards the driver, radars, accelerometers and angular rate sensors, and a global positioning system (GPS). In addition, data from the Control Area Network (CAN) bus in the vehicle was also collected, including information such as turn indicator usage and ego vehicle speed. Furthermore, the longitudinal and lateral positions, the speed of surrounding vehicles, and traffic density were also collected. In this study, we used data from the ego vehicle described below to determine the start and completion of a lane change manoeuvre.

Determining the start and completion of a lane change from time-series data can be challenging since the measured values are highly sensitive to the definitions selected by authors [33]. As a result, researchers often use various methods to identify lane change start and completion in their data. What can and cannot be done primarily depends on which data is available [34]. The thresholds of lateral distance to a lane marking and the lateral velocity of the vehicle are common variables used to determine a lane change’s start and completion point. This approach has been used on data collected from vehicle-mounted cameras [35] and data from high-rise building-mounted cameras [36]. However, neither study provided quantitative thresholds

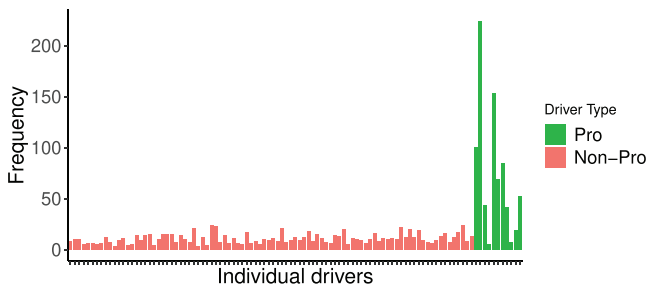
for lateral position or velocity. Thresholds provide objective, quantifiable criteria to determine lane change start and end-points. For example [33], used a  $-0.2$  m/s and [37], used a  $0.33$  m/s lateral velocity threshold to determine the start and end of a lane change.

In this study, we first filtered lane change scenarios using L3Pilot’s scenario definition, which only considered the distance to lane markings [31]. However, we used thresholds of lateral velocity and distance to lane markings as criteria to determine lane changes’ start and completion points. To explore the appropriate thresholds, we first analyzed the driver’s lateral velocity and position within the lane in a free-driving (the subject vehicle following its path without any influence of other vehicles) scenario. We found that in 90% of free driving time, the lateral speed did not exceed  $0.17$  m/s both driving towards the left and the right. Moreover, the lateral position within the lane was found to be centered in the middle of the lane and to be within  $0.42$  m to the right and  $0.36$  m to the left 90% of the time.

Assuming that the lateral velocity at the start and completion of the lane change is zero [10], or using a smaller threshold, may result in underestimating the start and completion of lane changes. On the other hand, using a larger threshold could result in identifying the start or completion of lane change after the vehicle has already crossed the lane boundary. Therefore, we tested different thresholds of lateral velocity to determine the start and completion of a lane change. We selected five thresholds from  $0.10$  m/s to  $0.3$  m/s (with an increase of  $0.05$  m/s) in the lane change scenarios. We also verified random cases for each chosen threshold by watching a front-facing video of the lane-changing vehicle to see if the lane change occurred at a particular threshold. We found that vehicles already crossed the lane boundary while using a threshold of  $0.20$  m/s to  $0.3$  m/s. Finally, we decided to use a threshold of  $0.15$  m/s to determine the start of a lane change. We selected this threshold to strike a balance between instances where a vehicle may have already crossed lanes and instances where a vehicle is fluctuating within its current lane. To cover the situations where a lane change might happen slowly without reaching the lateral velocity criteria, we also included cases with a lateral motion that got closer than  $5$  cm from the lane marking at a start of a lane change. The lateral distance was measured between the side of the vehicle and lane markings of the target lane. By selecting a smaller value such as  $5$  cm, we ensured that the vehicle was getting closer to the target lane marking. The target lane markings were determined from the direction of lane change. It is important to note that the study’s focus was exclusively on the start of lane changes. Therefore, all variables used in the model were extracted at the start of the lane change, as defined by the  $0.15$  m/s lateral speed, complemented with the  $5$  cm-from-lane-marking criterion.

### 2.1 | Descriptive data

A total of 1791 lane change cases for 103 drivers were extracted from the dataset, having met either of the previously described thresholds. The driver’s age ranged from 23 to 64 years old,



**FIGURE 1** The frequency of LC cases per driver.

with an average of 40.2 years old. The drivers consisted of 71 males and 32 females. Among these drivers, 69% had more than 10 years of driving experience, while the remaining drivers had shorter or non-specified driving experience.

As shown in Figure 1, there is an imbalance in the number of lane changes per driver. Of the 103 drivers in the dataset, only eleven (about 11%) were professional drivers. Despite this, they accounted for nearly 40% of the lane changes. This was due to the fact that professional drivers drove more and thus generated more driving data compared to non-professional drivers who typically drove for only 1 h. As mentioned earlier, the exclusion of trips where the automated driving function was activated has led to an imbalance among the trips of professional drivers. Furthermore, although the non-professional drivers drove for a similar duration, they exhibited variation in their driving patterns. This variation can be attributed to the fewer situations where a lane change was required. To avoid losing valuable insights, we decided to include data from both professional and non-professional drivers. However, we conducted an additional supplementary analysis using individual datasets for each group to ensure that the inclusion of both groups did not introduce any biases or obscure unique behavioural patterns specific to each group. Please refer to the Appendix for more details.

Figure 2 illustrates a typical lane change scenario, which depicts the different vehicles involved. The vehicles shown in red (rear and lag vehicles) potentially impact the turn signal usage behaviour of the lane-changing vehicles (shown in blue). This is because the turn signal is primarily used to communicate to the rear and lag vehicles that a lane change is imminent.

Overall turn signal usage was 92.8%, which is higher than reported in previous literature. However, further examination revealed a discrepancy in turn signal usage, specifically in relation to lane change start. In 59.9% of cases, a turn signal was used before starting a lane change, while it was only used after the start of a lane change in 32.9% of instances. Furthermore, in 7.2% of cases, no turn signal was used at all.

The following variables were taken into consideration for each lane change case: speed (km/h) of the ego vehicle, direction of lane change (right or left), presence of a rear vehicle (yes or no), rear gap (seconds), presence of a lag vehicle (yes or no), lag gap (seconds), and traffic density (vehicles/km/lane) and driver type (pro and no pro). The variables were selected because of their potential effect on turn signal usage. For instance, it was demonstrated in [15] that turn signal usage var-

ied depending on the direction of the lane change. Research has also shown that surrounding vehicles and their respective gaps impact the overall lane change process [38], although not directly on turn signal usage.

### 3 | BAYESIAN HIERARCHICAL MODEL

The realistic driving data, such as the one we used, are often large, noisy, sparse, and unbalanced [29]. Furthermore, it typically has a hierarchical or multilevel structure, where observations are nested within participants. To account for the variation, hierarchical models such as mixed effect multinomial Logit models are typically used in the traditional approach [39]. Nonetheless, hierarchical models based on the traditional approach have some limitations that can be addressed by employing a Bayesian approach. Bayesian data analysis provides a more intuitive interpretation of results using probabilities as opposed to  $p$ -values in traditional approach. The Bayesian approach can handle more complex and non-linear relationships, as well as non-normally distributed data [40].

At its fundamental level, the Bayesian approach uses Bayes' theorem to update prior beliefs or knowledge with observed data to make inferences about unknown parameters [41]. Bayes' theorem is the fundamental rule that combines prior probability and likelihood to compute the posterior probability. The prior probability represents the initial belief about the parameters or hypotheses before observing any data. In contrast, the posterior probability is the updated belief about the parameters after observing the data. For an in-depth introduction to Bayesian data analysis, readers are referred to [42].

Bayesian hierarchical (or multilevel) models (BHM) are a class of statistical models that allow for the modelling of complex relationships within hierarchical or nested data structures [43]. In our case, the data for lane change measurements are obtained from various drivers. This data has a multilevel structure, which makes BHM an advantageous statistical modelling approach to capture the intricate relationships within the nested data structure. Although not unique to BHM, a key feature of BHM is that they allow for the estimation of both fixed and random effects, enabling the decomposition of variance in the response variable into components attributable to different levels of the hierarchy.

The categorical response variable signal has three levels: before, after, and no. The level "before" is chosen as the baseline. The model will estimate the person-specific effects of the predictors and the random intercept for each driver. In summary, the model includes fixed effects for speed, direction, rear vehicle, rear gap, lag vehicle, lag gap, traffic density, driver type, and random intercepts for Driver ID. The random intercept allows us to account for any variation across different drivers. The hierarchy of the model is explained below.

- **Random effects:** At this level, the focus is on the individual drivers represented by the variable Driver ID. The random intercept ( $b_{0j}$ ) accounts for the differences in the baseline of the response variable (signal) across different drivers. This allows the model to capture variations between drivers, which

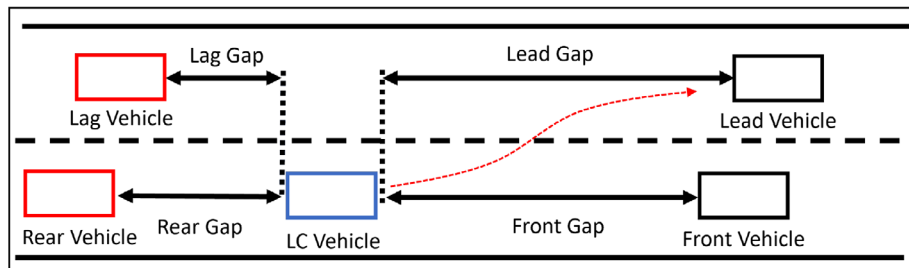


FIGURE 2 A typical LC scenario.

can help understand individual drivers' unique characteristics and behaviour.

- Fixed effects: At this level of the model, the response variable (Turn signal usage) is modelled as a function of the fixed effects, which include the predictors' (speed, direction, rear veh, rear gap, lag veh, lag gap, traffic density, and driver type). These fixed effects represent the average effects of each predictor on the response variable across all drivers. By including these fixed effects, the model captures the general relationship between the predictors and the response variable, allowing for an understanding of the overall trends in the data.

In this study, the probability of observing each level of the outcome variable (Signal) is modelled using Equations (1) and (2)

$$\text{logit}\left(\frac{P(\text{signal} = \text{after})}{P(\text{signal} = \text{before})}\right) = \alpha_{\text{after}} + \beta_{1,\text{after}} \times X + b_{0j,\text{after}} \quad (1)$$

$$\text{logit}\left(\frac{P(\text{signal} = \text{no})}{P(\text{signal} = \text{before})}\right) = \alpha_{\text{no}} + \beta_{1,\text{no}} \times X + b_{0j,\text{no}} \quad (2)$$

Here,  $X$  represents the vector of predictor variables,  $\alpha_{\text{after}}$  and  $\alpha_{\text{no}}$  are the intercepts for the "after" and "no" categories, respectively.  $\beta_{1,\text{after}}$  and  $\beta_{1,\text{no}}$  represent the fixed effect coefficients for the "after" and "no" categories, respectively, corresponding to the vector  $X$  of predictor variables. In these equations,  $b_{0j,\text{after}}$  and  $b_{0j,\text{no}}$  represent the random intercepts for each driver.

Integrating prior knowledge into the estimation process is a key advantage of Bayesian approach [44]. A prior is a probability distribution that expresses the degree of belief or uncertainty about the value of a parameter of interest before any data is observed. The prior distribution is used as a starting point to compute the posterior distribution, which represents the updated degree of belief or uncertainty about the parameter after observing the data. Nevertheless, this feature is also a source of common criticism, as there is no unique way of choosing a prior distribution, and prior specification can substantially influence the model outcomes [44, 45]. Without prior knowledge, often weakly informative or non-informative priors are assumed [42]. Weakly informative priors are sometimes used when some prior information about the parameters (e.g. the

direction of effect) is available. Still, it is not strong enough to justify a prior solid distribution. Conversely, non-informative priors are often used when no prior information is available or the analyst wants the data to determine the posterior distribution completely.

The BHM approach allows the use of priors while reducing the model's sensitivity to the choice of the prior distribution [45]. In BHM, the model parameters are organized into multiple levels or layers, each with its own prior distribution. In BHM, we incorporate additional information or structure into the model, which can help reduce the influence of the prior distribution on the resulting inference.

In our study, we used non-informative priors with normal distribution  $\mathcal{N}(0, 1 \times 10^4)$  for the fixed effects. This decision was made because there is no credible information available regarding the size or direction of the effects of different factors on turn signal usage. By using noninformative priors, we allow the data to influence the estimates most while still incorporating the prior information in the Bayesian framework. The priors for the random effects, such as for the standard deviations of the "after" and "no" parameters within each Driver ID group, are specified as half-normal  $\text{Half-normal}(0, 1 \times 10^4)$  distributions. Half-normal distributions are suitable for random effects because they are non-negative and non-informative [46]. They have minimal influence on the results while providing some regularization to improve convergence and sampling efficiency.

We also conducted a sensitivity analysis by comparing six sets of priors with different means and variance, and then analyzed their respective posterior distributions. The goal was to examine how the alternative prior specifications have affected the posterior distributions of the model parameters. If the posterior distributions are similar for different prior distributions, it implies that the posterior distribution is not sensitive to different priors [44]. We conducted the prior sensitivity analysis for each independent variable. However, to keep it concise, only the posterior distribution comparison graphs for two specific independent variables are provided in the Appendix. The plots indicate that the posterior distributions overlap and remain stable, suggesting low sensitivity to the selection of prior distributions [44].

The model is implemented using brms [47] package of R programming language [48], which provides an interface to fit Bayesian models using Stan. Stan is an open-source state-of-the-art platform for statistical modelling and high-performance statistical computation [49]. The model uses

**TABLE 1** Random effects for the hierarchical model.

Random effects	Estimate	Est.Error	l-95% HDI	u-95% HDI	Rhat	Bulk_ESS	Tail_ESS
<b>Standard deviation</b>							
<b>After</b>							
Intercept	0.36	0.11	0.13	0.57	1.00	2554	2557
<b>No</b>							
Intercept	2.69	0.55	1.58	4.19	1.00	2685	4289

**TABLE 2** Fixed effects for the hierarchical model.

Predictors	Estimate	Est.Error	l-95% HDI	u-95% HDI	Rhat	Bulk_ESS	Tail_ESS
<b>After</b>							
Intercept	-1.67	0.51	-2.68	-0.7	1.00	12,851	8654
Speed	0.02	0.01	0.01	0.04	1.00	12,714	8803
Direction (right) <sup>a</sup>	-0.81	0.11	-1.03	-0.59	1.00	17,352	7662
Rear vehicle (yes) <sup>b</sup>	-0.37	0.18	-0.74	-0.02	1.00	9408	7662
Rear gap	0.19	0.09	0.02	0.37	1.00	9111	7480
Lag vehicle (yes) <sup>b</sup>	-0.23	0.18	-0.59	0.13	1.00	9104	7691
Lag gap	0.07	0.09	-0.12	0.24	1.00	9031	7608
Traffic density	-0.02	0.01	-0.04	0.00	1.00	15,617	7884
Driver type (pro) <sup>c</sup>	0.30	0.18	-0.04	0.65	1.00	5447	6251
<b>No</b>							
Intercept	-5.92	1.34	-8.60	-3.41	1.00	5553	5741
Speed	0.01	0.01	-0.01	0.03	1.00	11,882	7803
Direction (right) <sup>a</sup>	0.02	0.22	-0.40	0.47	1.00	15,624	7487
Rear vehicle (yes) <sup>b</sup>	-0.44	0.38	-1.16	0.32	1.00	9182	7773
Rear gap	0.26	0.18	-0.10	0.60	1.00	11,790	8089
Lag vehicle (yes) <sup>b</sup>	-1.10	0.35	-1.83	-0.46	1.00	8098	7838
Lag gap	0.48	0.14	0.20	0.75	1.00	7283	7006
Traffic density	-0.07	0.03	-0.12	-0.02	1.00	16,416	7826
Driver type (pro) <sup>c</sup>	2.87	1.19	0.69	5.34	1.00	2647	4003

<sup>a</sup>Left is the reference level.

<sup>b</sup>No is the reference level.

<sup>c</sup>Non-pro is the reference level.

Markov Chain Monte Carlo (MCMC) sampling with four separate chains, where each chain comprises 10,000 iterations, including an initial 5000 burn-in iterations. During the burn-in phase, the sampler adapts and fine-tunes its parameters.

## 4 | RESULTS

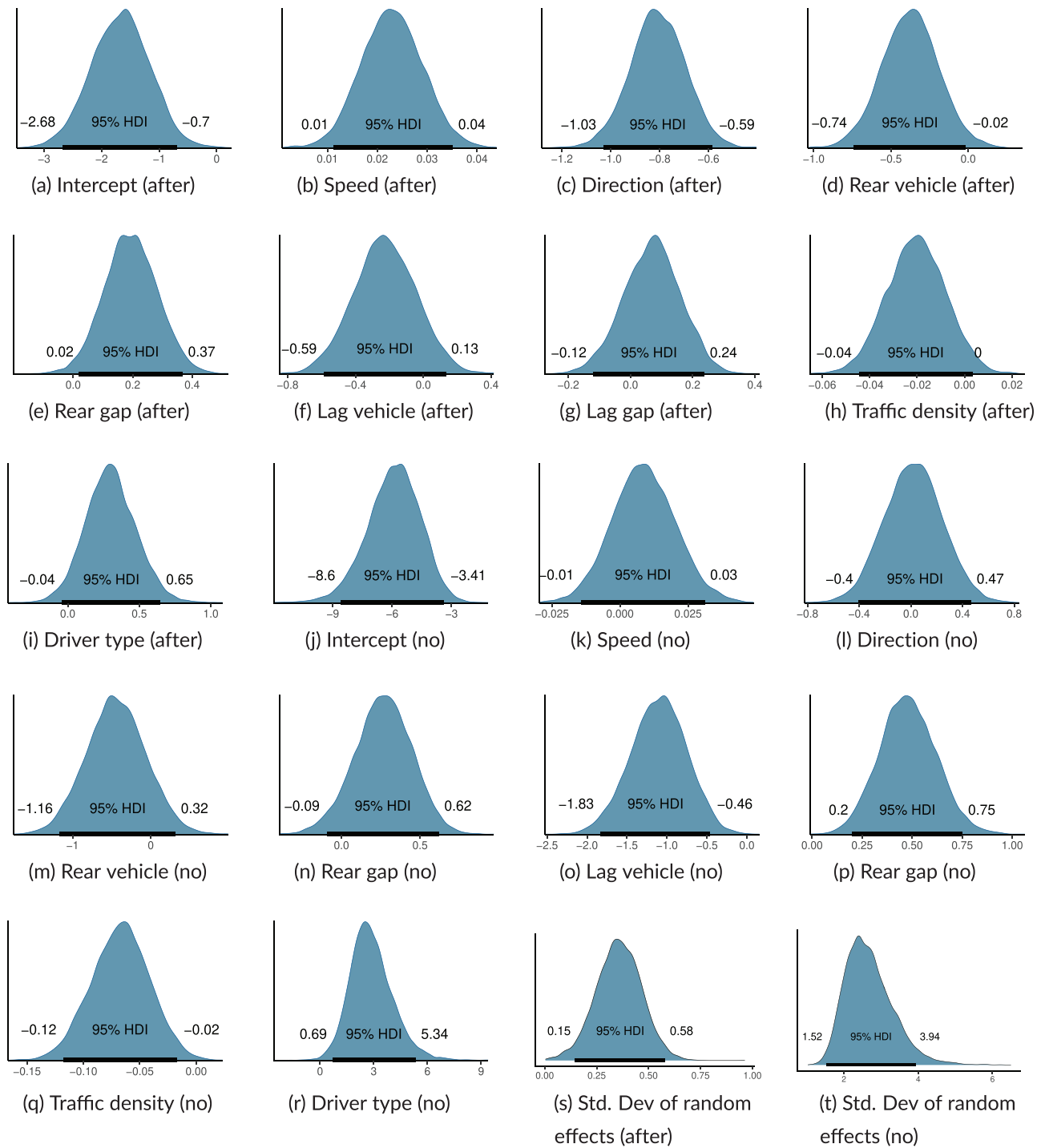
Figure 3 displays the posterior distribution for the fixed effects and the random effects parameters. These plots display the probability distribution of a parameter given the observed data and prior information. The plots provide a visual representation of uncertainty in the parameter estimates (see Tables 1 and 2), central tendency (e.g. mean), and spread. The uncertainty in the estimates is characterized by the width of the highest pos-

terior density interval (HDI). The HDI represents a range of values for the parameter, covering a certain percentage of the posterior distribution. These values indicate the most credible points within this distribution. A wider HDI indicates greater uncertainty, while a narrower HDI indicates more precision in the estimate [50]. If the HDI includes zero, it indicates that the true value of the parameter might not be significantly different from zero, given the data and the model [50]. The following sections provide a comprehensive analysis of the results.

### 4.1 | Random effects

Table 1 shows model results for random effects for the intercept Driver ID. A random intercept is an extension of the fixed





**FIGURE 3** The density plot display the distribution of samples drawn from the posterior distribution for the fixed and random effects. The x-axis represents the range of possible values of the parameter of interest, which is the regression coefficient for the predictor variable. The y-axis shows the density of the posterior distribution for each coefficient value. The density represents the relative frequency of the coefficient occurring in the posterior distribution. Each plot is annotated with the highest posterior density interval (horizontal at the bottom). The width of the HDI reflects the uncertainty in the parameter estimate.

intercept in a regression model that allows the intercept to vary across different levels of a grouping factor. In this case, we have a random intercept for each driver (Driver ID). This means that each driver has their baseline probability of the response variable (Signal) for the after and no categories after accounting for the fixed effects. The random intercepts capture the variation in

the baseline probabilities across drivers, which is not explained by the fixed effects (predictors).

The column “Estimate” represents the variability in the intercepts and slopes among different levels of the grouping factor (Driver ID in this case). The “Est.Error” measures the uncertainty or variability in the estimate. A larger standard error

indicates a greater degree of uncertainty in the estimate. The “l-95%” HDI and “u-95% HDI” is 95%, the range within which we can be 95% confident that the true parameter value lies. The “Rhat” is a convergence diagnostic used to assess whether the MCMC chains have converged to the target distribution. A Rhat value close to 1 indicates good convergence, and a value less than or equal to 1.1 is considered acceptable. Bulk ESS and Tail ESS represent the effective sample size (ESS) measures. The ESS estimates the number of independent draws from the target distribution to which the MCMC sample is equivalent. A higher ESS indicates that the MCMC sample provides more information about the target distribution.

The standard deviation of the random intercept for the level “after” suggests some variability in the “after” logits across different drivers. In other words, the likelihood of using a turn signal after starting a lane change varies across drivers when considering the average effect of all other factors. The standard deviation of the random intercepts for the “no” level suggests substantial variability in the “no” logits across different drivers. In other words, the likelihood of not using a turn signal while changing lanes varies across drivers when considering the average effect of all other factors.

## 4.2 | Fixed effects

In Table 2, estimates for the fixed effects in the “after” and “no” compared to the response level “before” of the response variable are presented. The sign of the coefficient of a predictor determines its relation with the outcome variable. For a continuous predictor, the estimates represent the change in log odds associated with a one-unit increase in the respective predictor, holding all other predictors constant. For categorical predictors, the estimates represent the change in log odds associated with a change from the reference category to the respective category, holding all other predictors constant.

The speed of the lane-changing vehicle is positively associated with the outcome variable level “after”. This suggests that the probability of using a turn signal after starting a lane change slightly increases with the per unit (km/h) increase in the speed of the subject vehicle. The speed of the lane-changing vehicle is also positively associated with the outcome variable level “no”. This indicates an increased probability of not using at all before starting a lane change with a per unit increase in the speed of a lane-changing vehicle. However, the HDI for the level “no” includes a zero indicating that the value of the parameter might not be significantly different from zero. Overall these results indicate that high speeds are maybe the potential motivation behind not using the turn signal properly while changing lanes.

The direction of lane change is negatively associated with the outcome variable level “after”. This indicates the probability of using a turn signal after starting a lane change decreases while making a lane change to the right compared to the left. This could be because rightward lane changes frequently occur after overtaking slower vehicles and returning to the initial lane, or because rightward lane changes (in right-hand traffic) are often mandatory. In both cases, drivers may be more likely to use turn signals correctly to inform surrounding traffic. Conversely,

the direction of the lane change has a positive association with the “no” outcome variable level, which seems to contradict the findings for the “after” level. However, the HDI includes zero, indicating that the parameter estimate may not be significantly different from zero. The discrepancy in the number of observations for each outcome variable level (after and no) could also contribute to the contradictory findings and uncertainty regarding the impact of lane change direction on the “no” outcome variable level.

Having a rear vehicle in the current lane is negatively associated with the outcome variable level “after”. This indicates that the probability of using a turn signal after starting a lane change decreases when a rear vehicle is in the current lane. However, a rear gap is positively associated with the outcome variable level “after”. This indicates that even if there is a rear vehicle, the probability of using a turn signal after starting a lane change increase with one unit increase in the rear gap. This is generally the case in real traffic where surrounding vehicles affect the lane change behaviour (e.g. as described in [38]). The association between the rear vehicle and the rear gap with outcome variable level “no” is the same as level “after”. However, the HDI for both predictors includes zero, indicating that the parameter estimate may not be significantly different from zero.

The presence of a lag vehicle shows a negative association with both “after” and “no” outcome variable levels, whereas the lag gap shows a positive association with these levels. Although these results align with those of rear vehicle and rear gap, the inclusion of zero in the confidence interval suggests uncertainty about the true impact on the outcome variable. The presence of rear and lag vehicles suggests increased overall traffic density. In our study, traffic density refers to the number of vehicles per kilometre per lane. The findings show a negative association between traffic density and both “after” and “no” outcome variable levels, suggesting that a one-unit increase in traffic density reduces the probability of turn signal usage being as “after” and “no.” However, the inclusion of zero in the HDI for the “after” level indicates that the parameter estimate may not be significantly different from zero.

Last but not least, driver type (pro) was positively associated with both “after” and “no” outcome variable levels. This suggests that the probability of using a turn signal after starting a lane change and not using a turn signal increases for professional drivers. However, the inclusion of zero in HDI for the “after” level and a very wide HDI for the “no” level indicates uncertainty about the true effect on the outcome variable. Furthermore, the discrepancy in the number of observations in the “no” level for the driver type could also contribute to these findings. Therefore, these findings should be interpreted with caution.

## 5 | DISCUSSION

This research examined the use of turn signals by Swedish (European) drivers during lane changes. Data for this study was collected on actual roads in Gothenburg, Sweden, observing drivers’ natural behaviour without interference from the experimenter. Out of the 103 participants, the majority were

non-professional drivers (everyday drivers), while approximately 10% were professional test drivers. The dataset included information on turn signal usage, the speed of the vehicle making a lane change, the direction in which the lane change occurred, the presence of and gaps between surrounding vehicles, as well as the overall traffic density. The turn signal usage was categorized into three levels: before starting a lane change, after starting a lane change, and not using the turn signal at all. The “before” level was selected as the baseline level due to its higher frequency and overall significance.

Our results showed that the turn signal usage in this study was very high, with approximately 93% of cases involving the use of turn signals. The results indicate higher compliance with turn signal usage than what has been reported in previous studies (e.g. ref. [2, 15, 16, 20, 33]). There may be several reasons for this high level of compliance. One possible reason could be that Swedish drivers have a higher compliance rate and a positive attitude towards adhering to traffic rules, as reported in previous studies [51, 52]. However, a substantial number (About 33%) of lane changes were observed where turn signals were used improperly. While there is no universally accepted definition of proper turn signal use in literature, it is generally understood to mean activating the turn signal before starting a lane change. The findings indicate comparable trends with the study of Lin and Bao [16], which revealed that in 70% of lane changes, a turn signal was used, but in half of those instances, the signal was activated after the lane change had already begun. Furthermore, our results indicated that in about 7% of the cases, a turn signal was not used at all. Overall, our results showed that only about 60% of the lane changes started with proper turn signal usage.

We further examined factors that impact turn signal usage through the application of Bayesian hierarchical modelling. The model consisted of fixed effects for variables such as speed, direction, rear vehicle, rear gap, lag vehicle, lag gap, traffic density, and driver type. Additionally, it included a random effect for Driver ID. The random effect for the Driver ID allowed to capture the variability in turn signal usage between different drivers. The results show that a range of factors influences turn signal usage, each with distinct levels of uncertainty. Highlighting the uncertainty in the findings not only acknowledges the fundamental principles of Bayesian analysis but also promotes a nuanced comprehension of the relationships and insights obtained from the data [50].

The results for the random effects indicated that the probability of using a turn signal after initiating a lane change is not consistent across drivers when considering the average effect of all other factors included in the model. Similarly, the probability of not using a turn signal while changing lanes is also inconsistent across drivers when considering the average effect of all other factors in the model. This variability suggests that individual driver behaviour and preferences may play a role in the likelihood of using a turn signal during lane change. Future research should investigate the factors contributing to individual differences in turn signal usage during lane changes. These factors may include but are not limited to, driving habits, personality traits, or other unmeasured factors.

The fixed effects in the model capture the overall relationship between the predictors and the response variable, representing the average effects of each predictor on the response variable across all drivers. The speed of the lane-changing vehicle was positively associated with both levels of the outcome variable. That means, with an increase in speed, the probability of either not using a turn signal or using it after starting a lane change also increased. A plausible reason for this could be that drivers tend to be less worried about vehicles in the intended lane when driving at faster speeds.

The direction in which a vehicle moves during a lane change can affect the overall manoeuvre. The direction of the lane change is typically associated with the underlying motivation for making the change. For example drivers may choose to make discretionary lane changes to the left (in a right-hand traffic scenario) to gain a speed advantage. Lane changes to the right typically occur due to mandatory reasons, such as taking an exit. The motivation behind a lane change subsequently influences the characteristics of the manoeuvre [20]. Our findings indicate that when drivers make lane changes to the right, they are less likely to use their turn signal after initiating the lane change. Conversely, the direction of the lane change has a positive association with the outcome variable level “no”. This observation is consistent with [15], which reported that during right lane changes in the US, drivers might not deem it necessary to signal their intentions after passing a slower lead vehicle. Please note that the results for the outcome variable “no” should be interpreted cautiously due to the inclusion of zero in the HDI.

According to prior research, the presence of vehicles in the current and target lane impacts lane change behaviour [38]. A slow vehicle in the current lane may prompt drivers to switch lanes in order to increase their speed. Similarly, the rear vehicle in the current lane and lag and lead vehicle in the target lane have an impact on overall lane change (e.g. gap acceptance or lane change duration). Similarly, the presence of a rear vehicle in the current lane, as well as lag and lead vehicles in the target lane, can influence the overall lane change process (e.g. gap acceptance or lane change duration). Our findings indicate that the probability of using a turn signal after starting a lane change, or not using it at all, decreases when there are rear and lag vehicles present. This is also reflected by the traffic density, which suggests that an increase in traffic density reduces the probability of using a turn signal after starting a lane change or not using it at all. However, as the rear gap and lag increase, the probability of using a turn signal after starting a lane change or not using it at all also increases. In other words, the impact of surrounding vehicles decreases with larger gap sizes. Future research should explore the specific gap at which the influence of the gap on turn signal usage during lane changes becomes negligible.

In our study, the main difference between driver types was their training and driving experience. Professional drivers were trained to operate prototype vehicles for testing specific features, while they had more experience behind the wheel than non-professional drivers. However, due to the imbalanced data and the discrepancy in the number of observations for each outcome variable level, the results concerning the driver type should be interpreted cautiously. Although a comparison

between professional and non-professional drivers was not a focus of this study, we also independently performed individual BHM analyses of the two datasets. The comparison results are provided in the Appendix. The results are provided in terms of posterior distribution density plots for both random and fixed effects.

## 5.1 | Implications

Our findings showed that human drivers often do not use turn signals properly during lane changes. These findings carry significant implications, which we explore in the following sections.

The AV technology's success depends on its mass adoption and smooth integration into existing traffic. Traffic is regarded as a complex social system, where road users make decisions based not only on traffic rules but also in response to the actions and anticipated actions of others [53–55]. One challenge that AVs will face in mixed traffic is communicating their intent and predicting/comprehending the intentions of human drivers. The use of turn signals is, therefore, a way through which human drivers announce their intention to surrounding drivers. Norman [56], regarded turn signals as the facial expression of vehicles. However, our findings indicated variability in human drivers' turn signal usage patterns. While variability itself highlights the unpredictability of human behaviour, AV algorithms may use this (probabilistic) information not just to predict when signals might be used but also to develop more conservative strategies when human-driven vehicles are likely to change lanes or turn without signalling. In situations where a turn signal is used, an AV can use this (probabilistic) information to improve conflict avoidance (e.g. modifying lane position as well as longitudinal position), to facilitate the generation of a safe trajectory before the actual lane change begins [57]. Our study identified several factors, such as the presence and gap to surrounding vehicles, that influence turn signal usage. An AV should be able to recognize these factors to anticipate potential lane changes by surrounding vehicles, even when they do not use turn signals, and subsequently prepare a safe trajectory in response.

Ponziani [2], pointed out that every instance where the use of a turn signal is neglected increases the risk of a crash. This study's observed patterns of turn signal usage during lane changes provide empirical evidence that can inform and shape policy decisions to improve safety in existing traffic systems. While turn signals are mandatory in most countries, clear rules on their proper use are lacking. For example a proper turn signal involves activating it for several seconds before changing lanes. While there may be no explicit regulations on the onset of turn signal activation by human drivers, the findings from our study can guide policymakers towards establishing clearer guidelines. Notably, the German Federal Ministry of Transport and Digital Infrastructure (BMDV) has recently set a precedent in this direction, mandating that AVs activate turn signals 3–5 s before starting a lane change [58]. In addition to general policy-making (which may be hard to enforce [2]), other measures can also be considered to increase the proper turn signal usage compliance. Two such measures – one

technology-based and one insurance-based – are discussed in the following sections.

Most new cars are equipped with advanced driver assistant systems such as lane departure warning (LDW) and lane keeping assistant (LKA). The LDW and LKA systems do not issue warnings if either the turn signal is activated or consistent steering movement (indicating a lane change) is detected. In the future, and with the right information with improved sensing, it may be possible for these systems to use algorithms to determine the context of driving (e.g. detection of lane change and its relation to surrounding traffic) and either automatically initiate turn signals, or prompt the driver to use the turn signal.

In addition to including more advanced technologies in our cars, using the capabilities of Pay-As-You-Drive (PAYD) insurance telematics, insurers may be able to play a pivotal role in enhancing turn signal usage compliance by incentivizing consistent and proper use of turn indicators during lane changes. For example insurers can give safety scores (linked to insurance premiums) based on proper turn signal usage, and provide the drivers with feedback if the turn signal usage is frequently ignored. However, further research is needed to investigate if such measures would be widely accepted and if PAYD is enough of an incentive to increase the turn signal.

## 5.2 | Limitations and future work

Finally, like any scientific study, this research also has limitations. The data were collected only in a specific city (Gothenburg) in Sweden. The road conditions, traffic laws, and driving behaviours in Sweden are unique to the country. Thus, the findings are most relevant to comparable settings in other Swedish cities and possibly neighbouring Scandinavian countries. Future studies should consider data from other parts of Europe (and other parts of the world) to better understand turn signal usage behaviour among different cultures. Additionally, the current study only focused on passenger cars. Given the current trajectory of research and development in the automotive sector, passenger cars are expected to be one of the first vehicles to widely adopt autonomous technology. As passenger cars move towards autonomy, it is crucial to gain insights also into their behaviour and patterns, across brands and over time (as the driving policies of automated vehicles evolve). As the proportion of automated vehicles and manual drivers on our roads increase, there may also be a shift in how human drivers use their turn indicators, depending on the driving policies of the automated vehicles. Further, the behaviour of large vehicles differs significantly from passenger cars. Therefore, future studies should also include heavy vehicles as their dynamics and behavioural patterns are distinct. The methodology used in our study could serve as the foundation for future studies.

The behaviour of changing lanes is also influenced by external factors such as the road type, time of day, and traffic environment. Our data was exclusively collected during the daytime on a single urban highway, known as the ring road in Gothenburg. These are the conditions (e.g. day time, urban highway) that are suitable for the (at least initial) widespread deployment of AVs.

Future studies should explore how different road types, times, and weather conditions affect turn signal usage. In this study, we only focused on how surrounding traffic affects turn signal usage. We showed the presence of lag, and rear vehicle impacts the turn signal usage. However, the type of vehicle could not be determined due to a high number of missing values for surrounding vehicles' lengths and widths. As a result, it would be valuable for future studies to take into account the types of surrounding vehicles present. We also found that the direction of lane change impacts the usage of turn signals. The direction could also be related to the motivation behind the lane change (e.g. discretionary or mandatory lane change), which in turn impacts the overall lane change process [59]. Therefore, future studies should consider investigating the differences in turn signal usage based on motivation. Finally, it is important to note that lane changes on divided highways with uninterrupted flow may differ from those that occur on approach lanes to intersections. In intersections, vehicles have more degrees of freedom in their movement [60]. Therefore, future studies should also consider studying turn signal usage behaviour during lane changes on the approach lanes of an intersection.

## 6 | CONCLUSION

The results of this study make a valuable contribution to the current body of literature by offering in-depth insights into the usage of turn signals during lane changes. Using turn signals while changing lanes has important effects on traffic safety and flow. Failing to use a turn signal while changing lanes can cause traffic disruption and increase the risk of collisions. However, our results show that turn signals are not used properly in a substantial portion of lane changes. Additionally, the use of turn signals is affected by various factors, including speed, the direction of lane change, and the presence of surrounding vehicles. Using turn signals as a direct means of communication is crucial for the future of mixed traffic, including both AVs and traditional vehicles. AVs are designed to follow strict rules, but they may face challenges in effectively communicating or negotiating right-of-way in certain situations [6, 61]. In this regard, a human driver's lane change without using a turn signal can cause problems for an AV. For instance, an AV may not be able to signal its intentions or force the driver to cancel a lane change. In such cases, the AV's reaction would likely be a sudden deceleration, potentially causing disruptions to traffic and discomfort for its occupants and other vehicles on the road.

In addition to the contribution to the existing literature, our findings likely have implications for general policy-making aiming to increase turn signal usage compliance, as well as possibly on the design of in-vehicle technologies (e.g. LDW and LKA), and PAYD insurance strategies. However, it is crucial to ensure that such measures are effective and accepted. Further research is therefore needed to cover more regions, driving cultures, road infrastructure and regulations, and vehicle types, as well as the acceptance of such safety measures.

## AUTHOR CONTRIBUTIONS

**Sarang Jokhio:** Conceptualization; methodology; formal analysis; data curation; writing—original draft; writing—review and editing; visualization. **Pierluigi Olleja:** Conceptualization; methodology; data curation; writing—review and editing. **Jonas Bårgman:** Writing—review and editing; supervision; project administration; funding acquisition. **Fei Yan:** Writing—review and editing. **Martin Baumann:** Writing—review and editing; supervision; project administration; funding acquisition.

## ACKNOWLEDGEMENTS

The study was carried out as part of the SHAPE-IT project, receiving funding from the European Commission's Horizon 2020 Framework Programme under the Marie Skłodowska-Curie Actions initiative (Grant agreement 860410). The authors thank Volvo Car Corporation. The authors also want to thank SAFER Vehicle and Traffic Safety Center at Chalmers, Gothenburg, Sweden, for providing us with the necessary facilities to extract data.

Open access funding enabled and organized by Projekt DEAL.

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

## DATA AVAILABILITY STATEMENT

The datasets generated and/or analyzed during the current study are not publicly available due to privacy concerns and restrictions imposed by the data collection source.

## ORCID

Sarang Jokhio  <https://orcid.org/0000-0003-1836-0664>

## REFERENCES

1. Kauffmann, N., Winkler, F., Naujoks, F., Vollrath, M.: What makes a cooperative driver? identifying parameters of implicit and explicit forms of communication in a lane change scenario. *Transp. Res. Part F: Traffic Psychol. Behav.* 58, 1031–1042 (2018)
2. Ponziani, R.: Turn signal usage rate results: a comprehensive field study of 12,000 observed turning vehicles (2012). Accessed 17 August 2022. <https://www.sae.org/content/2012-01-0261/>
3. StVO.: German road traffic regulations. (2021). §6. Bundesministerium für Verkehr und digitale Infrastruktur, Accessed 17 August 2022. <https://bmdv.bund.de/blaetterkatalog//catalogs/327056/pdf/complete.pdf>
4. DMV C.: Safe driving practices (2020). Accessed 17 August 2022. <https://www.dmv.ca.gov/portal/handbook/california-driver-handbook/introduction-to-driving/>
5. Stoll, T., Lanzer, M., Baumann, M.: Situational influencing factors on understanding cooperative actions in automated driving. *Transp. Res. Part F: Traffic Psychol. Behav.* 70, 223–234 (2020)
6. Färber, B.: Communication and communication problems between autonomous vehicles and human drivers. In: *Autonomous Driving*, pp. 125–144. Springer, Cham (2016)
7. Zheng, Z.: Recent developments and research needs in modeling lane changing. *Transp. Res. Part B: Methodol.* 60, 16–32 (2014)
8. Krajewski, R., Bock, J., Kloeker, L., Eckstein, L.: The highD dataset: a drone dataset of naturalistic vehicle trajectories on german highways for validation of highly automated driving systems. In: *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pp. 2118–2125. IEEE, Piscataway, NJ (2018)

9. Toledo, T., Koutsopoulos, H.N., Ben-Akiva, M.E.: Modeling integrated lane-changing behavior. *Transp. Res. Rec.* 1857(1), 30–38 (2003)
10. Li, Y., Li, L., Ni, D., Zhang, Y.: Comprehensive survival analysis of lane-changing duration. *Meas.: J. Int. Meas. Confed.* 182, 109707 (2021). doi:<https://doi.org/10.1016/j.measurement.2021.109707>
11. Ahmed, K., Ben-Akiva, M., Koutsopoulos, H., Mishalani, R.: Models of freeway lane changing and gap acceptance behavior. *Transp. Traffic Theory* 13, 501–515 (1996)
12. Goswami, V., Bham PhD, G.H.: Gap Acceptance Behavior in Mandatory Lane Changes Under Congested and Uncongested Traffic on a Multi-lane Freeway, number: 07-2919 (2007). Accessed 11 January 2023. <https://pubsindex.trb.org/view.aspx?id=802468>
13. Das, A., Khan, M.N., Ahmed, M.M.: Nonparametric multivariate adaptive regression splines models for investigating lane-changing gap acceptance behavior utilizing strategic highway research program 2 naturalistic driving data. *Transp. Res. Rec.* 2674(5), 223–238 (2020)
14. Tijerina, L., Garrott, W.R., Glecker, M., Stoltzfus, D., Parmer, E.: Van and passenger car driver eye glance behavior during the lane change decision phase. In: Proceedings of the 84th Annual Meeting of the Transportation Research Board. Transportation Research Board (1997)
15. Lee, S.E., Olsen, E.C.B., Wierwille, W.W.: A Comprehensive Examination of Naturalistic Lane-Changes: (733232011-001). American Psychological Association, Washington, D.C. (2004)
16. Lin, B.T., Bao, S.: The effect of turn signal onset on lateral performance measures when overtaking a lead vehicle-using naturalistic driving environment. In: Driving Assessment Conference, vol. 10, pp. 266–272. University of Iowa, Iowa (2019). Accessed 16 December 2022. [https://drivingassessment.uiowa.edu/sites/drivingassessment.uiowa.edu/files/da2019\\_42\\_brianlin\\_final.pdf](https://drivingassessment.uiowa.edu/sites/drivingassessment.uiowa.edu/files/da2019_42_brianlin_final.pdf)
17. Bezzina, D., Sayer, J.: Safety pilot model deployment: test conductor team report USDOT report no. DOT HS 812 171. (2015)
18. Fitch, G.M., Lee, S.E., Klauer, S., Hankey, J., Sudweeks, J., Dingus, T.: Analysis of lane-change crashes and near-crashes. Virginia Tech Transportation Institute, Blacksburg, VA (2009)
19. Dang, R., Zhang, F., Wang, J., Yi, S., Li, K.: Analysis of Chinese driver's lane change characteristic based on real vehicle tests in highway. In: 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013), pp. 1917–1922. IEEE, Piscataway, NJ (2013)
20. Wang, X., Yang, M., Hurwitz, D.: Analysis of cut-in behavior based on naturalistic driving data. *Accid. Anal. Prev.* 124, 127–137 (2019)
21. Russo, R.: Statistics for the Behavioural Sciences: An Introduction to Frequentist and Bayesian Approaches. Routledge, London (2020)
22. Depaoli, S., Rus, H.M., Clifton, J.P., van de Schoot, R., Tiemensma, J.: An introduction to Bayesian statistics in health psychology. *Health Psychol. Rev.* 11(3), 248–264 (2017)
23. Daziano, R.A., Miranda-Moreno, L., Heydari, S.: Computational Bayesian statistics in transportation modeling: from road safety analysis to discrete choice. *Transp. Rev.* 33(5), 570–592 (2013)
24. Huang, H., Abdel-Aty, M.: Multilevel data and bayesian analysis in traffic safety. *Accid. Anal. Prev.* 42(6), 1556–1565 (2010)
25. Xie, K., Wang, X., Huang, H., Chen, X.: Corridor-level signalized intersection safety analysis in Shanghai, China using Bayesian hierarchical models. *Accid. Anal. Prev.* 50, 25–33 (2013)
26. Zheng, L., Sayed, T., Essa, M.: Bayesian hierarchical modeling of the non-stationary traffic conflict extremes for crash estimation. *Anal. Methods Accid. Res.* 23, 100100 (2019)
27. Yu, R., Abdel-Aty, M.: Using hierarchical bayesian binary probit models to analyze crash injury severity on high speed facilities with real-time traffic data. *Accid. Anal. Prev.* 62, 161–167 (2014)
28. Deublein, M., Schubert, M., Adey, B.T., Köhler, J., Faber, M.H.: Prediction of road accidents: a Bayesian hierarchical approach. *Accid. Anal. Prev.* 51, 274–291 (2013)
29. Morando, A., Victor, T., Dozza, M.: A bayesian reference model for visual time-sharing behaviour in manual and automated naturalistic driving. *IEEE Trans. Intell. Transp. Syst.* 21(2), 803–814 (2019)
30. Schindler, R., Flannagan, C., Bálint, A., Piccinini, G.B.: Making a few talk for the many—modeling driver behavior using synthetic populations generated from experimental data. *Accid. Anal. Prev.* 162, 106331 (2021)
31. Penttinen, M., Rämä, P., Dotzauer, M., Hibberd, D., Innamaa, S., Louw, T., et al.: Deliverable D3.2 of L3Pilot: experimental procedure (2019). Accessed 22 November 2021. <https://www.l3pilot.eu/publications/deliverables.html>
32. Bundesagentur für Arbeit: Statistics of the german federal employment agency, tables, employees by occupation (kldb 2010) (quarterly figures). Nuremberg, (2018). Accessed 18 December 2022. [https://www.arbeitsagentur.de/datei/annual-report-2018\\_ba045416.pdf](https://www.arbeitsagentur.de/datei/annual-report-2018_ba045416.pdf)
33. Wang, Q., Li, Z., Li, L.: Investigation of discretionary lane-change characteristics using next-generation simulation data sets. *J. Intell. Transp. Syst.* 18(3), 246–253 (2014)
34. Xi, Y., Crisler, M.: A review of lane change definitions and identification methods. In: *Transportation Research Board 92nd Annual Meeting/Transportation Research Board*, pp. 13–1691. Transportation Research Board, Washington, D.C. (2013)
35. Olsen, E.C., Lee, S.E., Wierwille, W.W., Goodman, M.J.: Analysis of distribution, frequency, and duration of naturalistic lane changes. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 46, pp. 1789–1793. Sage Publications, Los Angeles, CA (2002)
36. Toledo, T., Zohar, D.: Modeling duration of lane changes. *Transp. Res. Rec.* 1999(1), 71–78 (2007)
37. Mullakkal-Babu, F.A., Wang, M., van Arem, B., Happee, R.: Empirics and models of fragmented lane changes. *IEEE Open J. Intell. Transp. Syst.* 1, 187–200 (2020)
38. Moridpour, S., Rose, G., Sarvi, M.: Effect of surrounding traffic characteristics on lane changing behavior. *J. Transp. Eng.* 136(11), 973–985 (2010)
39. Bhat, C.R., Gossen, R.: A mixed multinomial logit model analysis of week-end recreational episode type choice. *Transp. Res., Part B: Methodol.* 38(9), 767–787 (2004)
40. Kruschke, J.K., Liddell, T.M.: The Bayesian new statistics: hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychon. Bull. Rev.* 25, 178–206 (2018)
41. van de Schoot, R., Depaoli, S., King, R., Kramer, B., Märtens, K., Tadesse, M.G., et al.: Bayesian statistics and modelling. *Nat. Rev. Methods Primers* 1(1), 1 (2021)
42. Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B.: *Bayesian Data Analysis*. CRC Press, Boca Raton, FL (2013)
43. Gelman, A., Hill, J.: *Data Analysis Using Regression and Multi-level/Hierarchical Models*. Cambridge University Press, Cambridge (2006)
44. Depaoli, S., Winter, S.D., Visser, M.: The importance of prior sensitivity analysis in bayesian statistics: demonstrations using an interactive shiny app. *Front. Psychol.* 11, 608045 (2020)
45. Robert, C.P., et al.: *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*, vol. 2. Springer, Berlin, Heidelberg (2007)
46. Gelman, A.: Prior distributions for variance parameters in hierarchical models. *Bayesian Anal.* 1(3), 515–534 (2006). doi:<https://doi.org/10.1214/06-BA117A>
47. Bürkner, P.C.: brms: an R package for Bayesian multilevel models using stan. *J. Stat. Software* 80, 1–28 (2017)
48. R Core Team: A language and environment for statistical computing. Vienna, Austria, (2022). Accessed 10 April 2023. <https://www.R-project.org/>
49. Carpenter, B., Gelman, A., Hoffman, M.D., Lee, D., Goodrich, B., Betancourt, M., et al.: Stan: a probabilistic programming language. *J. Stat. Software* 76(1), 1–32 (2017)
50. Kruschke, J.K.: Bayesian estimation supersedes the *t* test. *J. Exp. Psychol.: General* 142(2), 573–603 (2013)
51. Warner, H.W., Özkan, T., Lajunen, T.: Cross-cultural differences in drivers' speed choice. *Accid. Anal. Prev.* 41(4), 816–819 (2009)
52. Sinclair, M.: Attitudes, norms and driving behaviour: A comparison of young drivers in south africa and sweden. *Transp. Res., Part F: Traffic Psychol. Behav.* 20, 170–181 (2013)

53. Wilde, G.J.: Social interaction patterns in driver behavior: an introductory review. *Hum. Factors* 18(5), 477–492 (1976)
54. Juhlin, O.: Traffic behaviour as social interaction-implications for the design of artificial drivers. Paper presented at the 6th world congress on intelligent transport systems (ITS), Toronto, Canada, 8–12 November 1999
55. Stoll, T., Strelau, N.R., Baumann, M.: Social interactions in traffic: the effect of external factors. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 62, pp. 97–101. Sage Publications, Los Angeles, CA (2018)
56. Norman, D.: *Turn Signals are the Facial Expressions of Automobiles*. Diversion Books, New York (2014)
57. Yoneda, K., Kuramoto, A., Sukanuma, N.: Convolutional neural network based vehicle turn signal recognition. In: *2017 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)*, pp. 204–205. IEEE, Piscataway, NJ (2017)
58. BMDV: Regulation on the approval and operation of motor vehicles with autonomous driving function in defined operating areas (autonomous vehicles approval and operation regulation - afgbv). (2022). [https://bmdv.bund.de/SharedDocs/DE/Anlage/K/presse/008-verordnung-automatisierte-autonome-fahrfunktion.pdf?\\_\\_blob=publicationFile](https://bmdv.bund.de/SharedDocs/DE/Anlage/K/presse/008-verordnung-automatisierte-autonome-fahrfunktion.pdf?__blob=publicationFile)
59. Vechione, M., Balal, E., Cheu, R.L.: Comparisons of mandatory and discretionary lane changing behavior on freeways. *Int. J. Transp. Sci. Technol.* 7(2), 124–136 (2018)
60. Zhao, J., Knoop, V.L., Wang, M.: Microscopic traffic modeling inside intersections: interactions between drivers. *Transp. Sci.* 57(1), 135–155 (2023)
61. Metz, D.: Developing policy for urban autonomous vehicles: impact on congestion. *Urban Sci.* 2(2), 33 (2018)

**How to cite this article:** Jokhio, S., Olleja, P., Bärghman, J., Yan, F., Baumann, M.: Exploring turn signal usage patterns in lane changes: A Bayesian hierarchical modelling analysis of realistic driving data. *IET Intell. Transp. Syst.* 1–16 (2023).  
<https://doi.org/10.1049/itr2.12457>

## APPENDIX

### Additional analysis

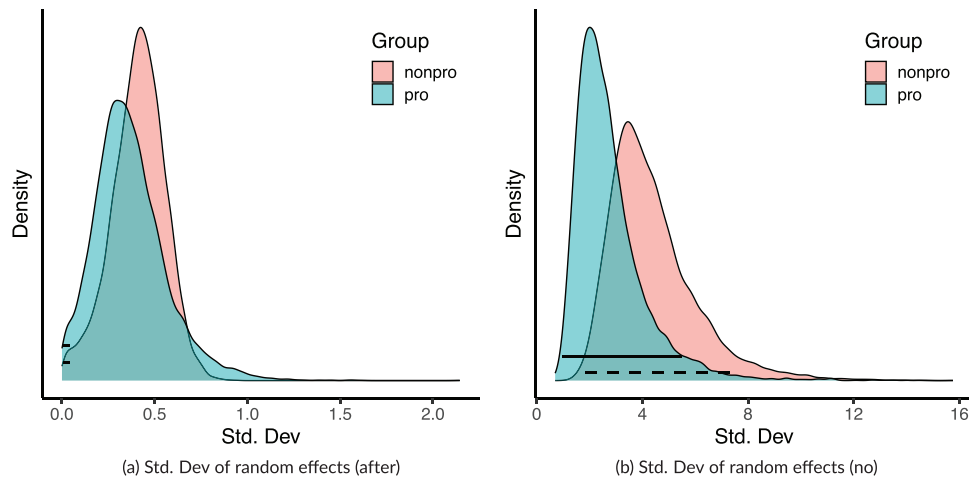
This section presents a comparison of the additional models that we ran. The priors and the parameters (e.g. a number of

chains and iterations) for each additional model were kept the same as in the original model. Please note that the number of observations and drivers in each dataset differs. This difference in sample size might affect the precision of the estimates and the width of the HDI intervals. Figure A.1 shows the posterior distribution for the random effects for both levels of the outcome variable (i.e. after and no). The extent to which both plots overlap can give us an idea of the difference between the two groups [50]. For example, Figure A.1(a) shows that overlap in posterior distribution for the outcome variable level “after” is lower than the level “no” in Figure A.1(b). In other words, professional drivers show less variability than non-professional drivers. However, further research is required considering a more balanced dataset.

Figure A.2(n) shows the posterior distribution for the fixed effects for both levels of the outcome variable (i.e. after and no). For the level “after” we can see a substantial variability only for the direction of lane change and traffic density as shown in Figure A.2(b,g). However, for the level “no” we can see a substantial variability in all the variables. This is because the all cases where a turn signal was not used belong to the professional driver group.

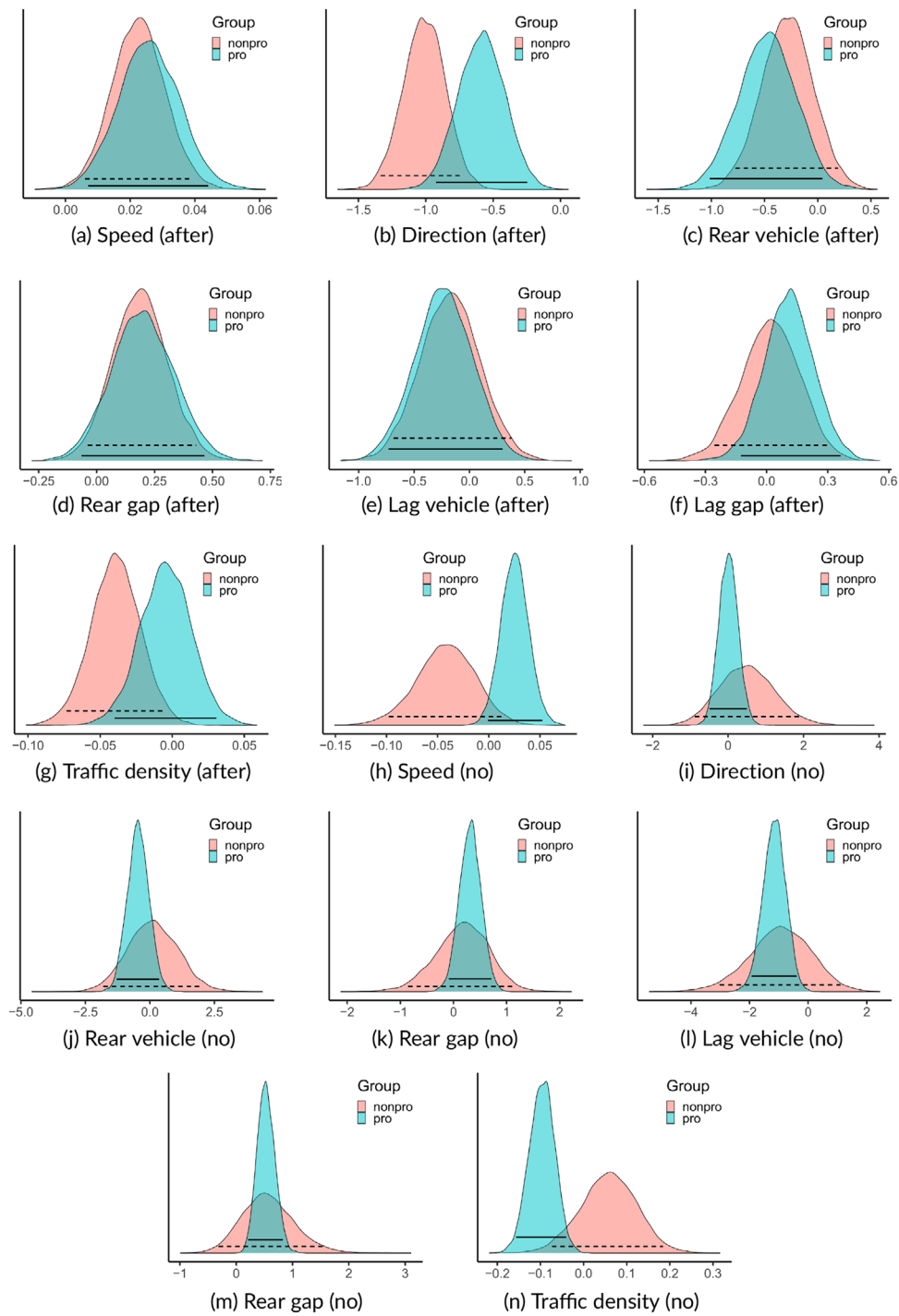
### Sensitivity analysis

This section shows the posterior distribution density plots comparing six different prior sets. The fixed effects were evaluated using six distinct priors: the original priors followed a normal distribution with a mean of 0 and a standard deviation of  $10^4$ , while the five alternative sets varied between means of 0 and  $-1$ , and standard deviations ranging from  $10^4$  to 1. For random effects, the original priors were based on a half-normal distribution with a mode of 0 (to ensure non-negative values) and a standard deviation of  $10^4$ , with two alternative sets having the same mode but standard deviations of 10 and 1, respectively. The posterior distributions plots in the Figure A.3 overlap and remain stable, suggesting low sensitivity to prior distributions.

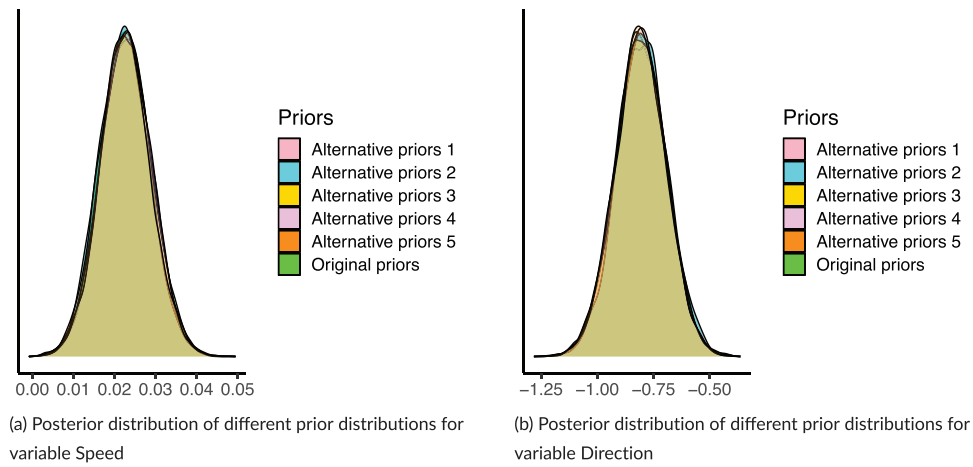


**FIGURE A.1** The density plot display the distribution of samples drawn from the posterior distribution for the random effects for both levels of outcome variable (i.e. after and no). The  $x$ -axis represents the range of possible values of the parameter of interest, which is the standard deviation. The  $y$ -axis shows the density of the posterior distribution for each coefficient value. The density represents the relative frequency of the coefficient occurring in the posterior distribution. Each plot is annotated with the highest posterior density interval (horizontal line at the bottom). The width of the HDI reflects the uncertainty in the parameter estimate.





**FIGURE A.2** The density plot display the distribution of samples drawn from the posterior distribution for the fixed effects for both levels of outcome variable (i.e. after and no). The  $x$ -axis represents the range of possible values of the parameter of interest, which is the regression coefficient for the predictor variable. The  $y$ -axis shows the density of the posterior distribution for each coefficient value. The density represents the relative frequency of the coefficient occurring in the posterior distribution. Each plot is annotated with the highest posterior density interval (horizontal line at the bottom). The width of the HDI reflects the uncertainty in the parameter estimate.



**FIGURE A.3** Comparative posterior distributions of six different priors. This visualization shows the differences in the posterior outcomes when the model is subjected to six varied prior specifications, highlighting the influence of prior beliefs on the posterior inferences.