Effect of road safety-conscious motorcycle taxi programs on riding behaviors and risk of road traffic crash among motorcycle taxi drivers in Kampala, Uganda

Kennedy M. Muni

A dissertation
submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

University of Washington

2019

Reading Committee:

Brandon Guthrie, Chair

Charles Mock

Philip Hurvitz

James Hughes

Program Authorized to Offer Degree:

Epidemiology

University of Washington

Abstract

Effect of road safety-conscious motorcycle taxi programs on riding behaviors and risk of road traffic crash among motorcycle taxi drivers in Kampala, Uganda

Kennedy M. Muni

Chair of the Supervisory Committee:

Brandon Guthrie

Departments of Global Health and Epidemiology

Background: *SafeBoda* is a private transportation company that started providing motorcycle taxi services in Kampala in 2015. Prior to the introduction of *SafeBoda*, motorcycle taxi services in Kampala were entirely provided by independent motorcycle taxi drivers with limited formal training on road safety and proclivity for risky riding behaviors (e.g., driving without protective gear and against traffic). *SafeBoda* introduced an Uber-like model of motorcycle taxi services where it provides road safety training, protective gear, and vehicle maintenance to its drivers. In addition, its drivers are expected to adhere to a code of conduct (e.g., respect for traffic regulations such as stopping at traffic stops). Penalties for breaking the code of conduct include re-training, suspension, or dismissal. We sought to determine whether helmet use, risk of road traffic crash (RTC), and riding behaviors differed between *SafeBoda* and regular (i.e., those not enrolled in *SafeBoda*) motorcycle taxi drivers in Kampala, Uganda.

Method: We collected demographic and behavioral data from *SafeBoda* and regular drivers using: a) computer-assisted personal interview (CAPI), where 400 drivers were asked about their riding behaviors (e.g., helmet and mobile phone use); b) roadside observation questionnaire, where riding behaviors were observed in 3000 boda-boda drivers and their passengers along major roads in Kampala; c) text-based/SMS and telephone interview questionnaires where occurrence of road traffic crash data was collected from 342 drivers every two months for a period of six months; and d) Global Positioning System (GPS) devices where movement patterns and riding data were collected from 60 drivers for a period of 24 hours. Baseline characteristics were compared between *SafeBoda* and regular drivers using chi-square and t-tests. In addition, we used Poisson and generalized estimating equation models with robust standard errors to model the effect of the *SafeBoda* program on helmet use and risk of road traffic crash respectively.

Results: Across all studies, a higher proportion of *SafeBoda* drivers than regular drivers engaged in safe riding behaviors. For instance, helmet use among *SafeBoda* compared to regular drivers was 21 percent points higher (95% CI: 0.15-0.27; p<0.001) based on the CAPI and 45 percent points higher (95% CI: 0.43-0.47; p<0.001) based on roadside observation. Furthermore, compared to regular drivers, *SafeBoda* drivers were more likely to report having a driver's license (66.3% vs 33.5%; p<0.001) and a reflective jacket (99.5% vs 50.5%; p<0.001) and were less likely to report driving against traffic (4% vs 45.7%; p<0.001) in the past 30 days. From the follow-up study of 342 drivers for 6 months, there were 85 crashes (31 in *SafeBoda* and 54 in regular drivers) that occurred during the follow-up. The majority of the 85 crashes (31 in *SafeBoda* vs 54 in regular drivers) involved either a collision with another motorcycle (27.1%) or a car (63.5%). Speeding (10.6%), faulty brakes (7.1%), and distracted driving (4.7%) were the most frequently reported causes of the crashes. Sixty-nine (81.2%) of the crashes resulted in injury to the driver and 56 (81.2%) of these driver injuries required a visit to a health facility. Of the injuries that required a visit to a health facility, 13 (23.2%) required in-patient care (admission). The median hospitalization time for injuries requiring in-patient care (as of the time of follow-up) was 3 days with a range from 1 day to 30 days. Over the six-month follow-up period, *SafeBoda* drivers were 39% less likely to be involved in a RTC than regular drivers after adjusting for age, possession of a driver's license, and education (RR: 0.61, 95% CI: 0.39-0.97, p=0.04). From the GPS study, we found that GPS devices are acceptable and feasible for measuring boda-boda driver movements without any major measurement (e.g., missing data and device failure) and logistical (e.g., installation and retrieval of the devices) issues. The GPS data showed that boda-boda drivers in the study made an average of 31 trips per day (SD = 10.4). The median trip was 3.5 km and lasted on average 9.0 minutes (range 3 minutes to 13.8 minutes). The mean farthest Euclidean distance from the driver's stage or taxi stand was 5.8 km. The mean speed on a trip was 22.5 km/h. Driving movements within the city did not seem to differ significantly between *SafeBoda* and regular drivers (they shared similar activity spaces). Even where there were differences, these seem to be quite modest.

Conclusion: The *SafeBoda* program is associated with increased safe riding behaviors and reudced risk of road traffic crash among motorcycle taxi drivers in Kampala. Therefore, the promotion and expansion of such programs may lead to a reduction in morbidity and mortality due to road injuries.

Table of contents

Acknowledgements

I would like to express my sincere appreciation to those who supported me throughout my training at the University of Washington. I am very grateful to my Doctoral Committee for their support and mentorship: Charlie Mock, Philip Hurvitz, Jim Hughes, and my chair and academic advisor Brandon Guthrie. I would also like to acknowledge the mentorship and support of Olive Kobusingye at Makerere University School of Public Health. I am thankful to Grace Magambo, Simon Okurut and Joseph Mugisha who provided logistical and data collection support during the course of my dissertation research. I thank the motorcycle taxi drivers whose participation made these studies possible. Lastly, I am thankful to my family and friends for their support during my years of training. This dissertation work was financially supported by the Fogarty International Center (Grant # R25 TW009345).

Dedication

To my parents Rose Juan Muni & Sadaraka Muni

Chapter 1: Introduction

Road traffic crashes kill 1.2 million and injure between 20 million and 50 million persons each year. Nearly all (95%) of this injury burden is borne by road users in low and middle income countries (LMICs).[1-3] Road traffic crashes (RTCs) involving motorized 2-3 wheelers (including motorcycles) account for 22% of these deaths.[1] Motorcycles are a popular mode of transportation in many LMICs because compared to other modes of transportation, they are relatively affordable, convenient for short distances, difficult terrain, and poor roads, and are efficient for navigating heavy vehicular traffic in many developing cities.[4-9] However, despite these benefits, motorcycle drivers are involved in a substantial number of RTCs and road traffic injuries (RTIs) in LMICs.[1-3,10] Africa has the world's highest rates of road traffic morbidity and mortality, with an estimated 65.2 RTIs and 26.6 deaths per 100000 persons.[3,11] Among motorcycle drivers and their passengers, the injury and death rates in Africa are 16.1 RTIs and 1.3 deaths per 100000 persons respectively.[3] In the past two decades, many studies have reported increasing involvement of motorcycle drivers and their passengers in RTCs and RTIs, as the number of motorcycles exploded in Africa.[3,6,9,12-14] One study in Nigeria reported an increase of 3000% in the number of patients with fractured limbs after the introduction of motorcycle taxis in Lagos, Nigeria.[15] Low helmet use (with estimates ranging from 0% to 89% among riders), low percent of riders with driving license, poor road conditions, poor enforcement of traffic laws, distracted driving behaviors (e.g., mobile phone use), and limited access to road safety trainings have been put forward as the primary drivers of the high incidents of RTC and RTI among motorcycle riders.[5,6,8,16-23] Unlike in high-income countries where mandatory training, licensing, and enforcement of traffic laws have ensured high compliance with use of protective gear[8,24,25], helmet use is low in Africa. Additionally, training is poor. For example, between 19%

and 80% of motorcycle drivers in Africa acquired their training informally from friends and peers.[18,23]

In Uganda, motorcycles are primarily used a form of taxi service and are known locally as boda-boda.[26-28] In the city of Kampala alone, the number of motorcycle taxi drivers is estimated to be between 50000 and 80000.[27] In line with patterns observed in other LMICs, motorcycle taxi drivers in Uganda are known to be involved in risky riding behaviors (e.g., speeding, unsafe overtaking, driving under the influence of drugs, and failure to wear protective gear) associated with increased risk of RTC and RTI. They are the second largest road user category involved in RTCs in Uganda. They were involved in 28.2% of RTCs and represented 23.5% of road traffic injuries in 2015.[10] According to the 2018 Road Safety Performance Review report, road fatalities involving motorcycle taxi drivers doubled between 2011 and 2016.[29]

In recent years, road safety-conscious companies such as *SafeBoda, UberBoda,* and *Taxify* have cropped up to formalize the motorcycle taxi sector in Uganda and to address the issues around poor riding behaviors through training and provision of protective gear.[30,31] *SafeBoda* is a private transportation company that takes an innovative approach to providing motorcycle taxi services using a trained community of drivers in an effort to increase safe riding behaviors and reduce crashes and injuries. The company provides multiphase road safety training, helmets, vehicle maintenance, and basic first responder training to its drivers. It also provides hairnets to passengers who are concerned about contracting skin diseases from a shared helmet.[30] In addition, *SafeBoda* drivers are expected to adhere to a code of conduct focused on safe driving and customer service. Initially, the training for the drivers was provided by the Uganda Police and the Uganda Red Cross Society. However, this has since been transitioned to a team of *SafeBoda* trainers (part of the *SafeBoda* academy) who were trained by the Global Road Safety Partnership and the Uganda Police and Red Cross. Newly recruited drivers are trained on traffic signs and symbols, traffic regulations, the *SafeBoda* code of conduct, emergency response, customer care, and how to use the *SafeBoda* app.[30,32,33] *SafeBoda* drivers undergo regular refresher trainings. As of January 2019, *SafeBoda* has over 8000 drivers in Kampala.[33] Other programs working in the same motorcycle taxi service space with *SafeBoda* include *UberBoda*, *Taxify*, and *Dial Jack*.

Despite the proliferation of these programs in Uganda, there is inconclusive evidence on whether road safety programs geared towards motorcycle drivers in LMICs are associated with increased safe riding behaviors and reduced incidents of RTC and injuries.[2,6,34-36] Moreover, no studies have been conducted in Uganda looking at the role of recent start-ups (e.g., *SafeBoda, Taxify, and UberBoda*) on road safety in the motorcycle taxi sector.

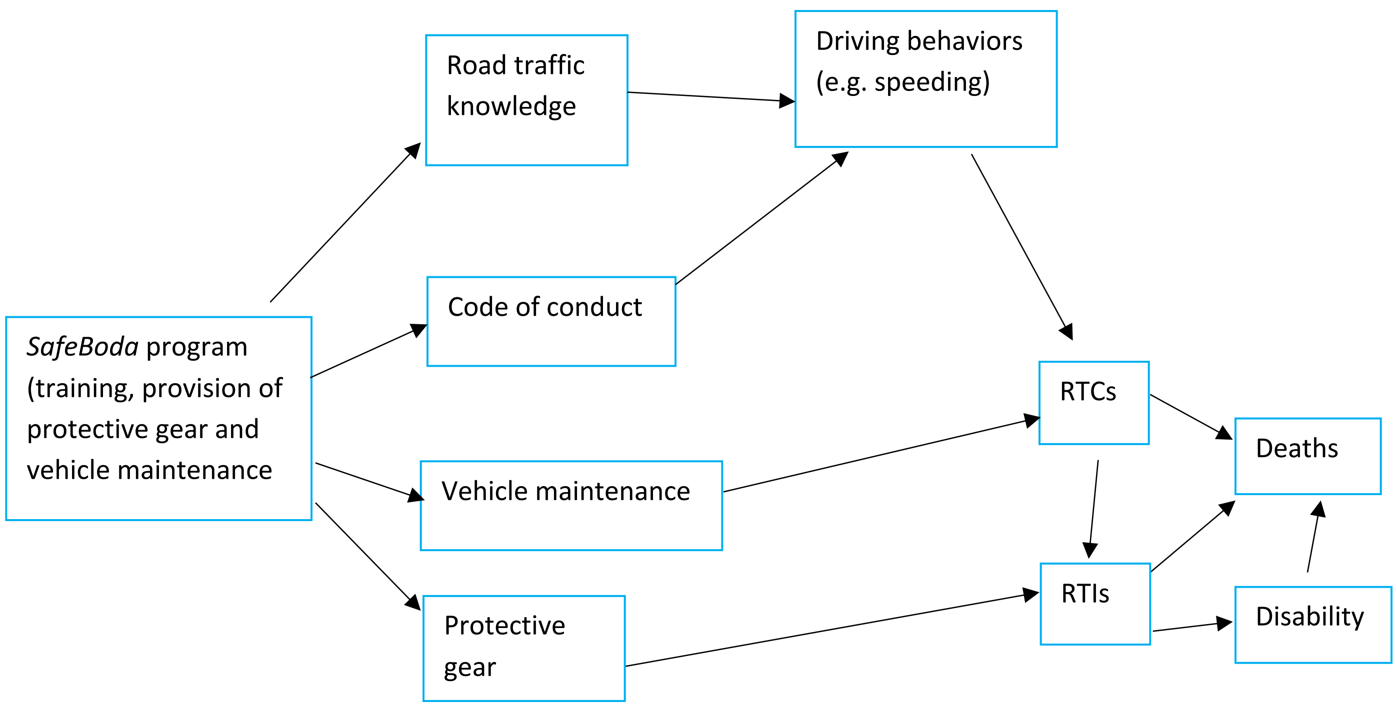
This dissertation set out to determine whether enrollment in the *SafeBoda* was associated with increased safe riding behaviors (e.g., use of helmet while driving) and reduced risk of road traffic crash (Figure 1). The dissertation is composed of three chapters as outlined below.

Chapter 2: We conducted three inter-related cross-sectional studies to determine whether a higher proportion of *SafeBoda* drivers and their passengers engaged in safe riding behaviors compared to regular drivers and their passengers. We collected demographic characteristics and riding behaviors from 400 motorcycle taxi drivers using a computer-assisted personal interview (CAPI). In addition, we observed helmet use in 80 participants from the CAPI study and in 3000 drivers and their passengers along purposively selected streets in Kampala.

Chapter 3: We conducted a cohort study where 342 drivers were followed up and assessed for occurrence of road traffic crash via text-based questionnaire and telephone interview every two months for a period of 6 months. In addition, we assessed whether road traffic knowledge (based on a score from common road traffic signs and laws) mediated the hypothesized relationship between the *SafeBoda* program and risk of RTC.

Chapter 4: We conducted a cross-sectional study to assess the feasibility of using global positioning system (GPS) devices to measure movement patterns and riding behaviors in a sample of 60 motorcycle taxi drivers drawn from the cohort study in Chapter 3. We also assessed in these movement patterns and driving behaviors differed between *SafeBoda* and regular drivers.

Figure 1.1. Conceptual model to visualize the relationships between the *SafeBoda* program, driving behaviors, and risk of road traffic crash



Chapter 2: Motorcycle taxi program increases safe riding behaviors among its drivers in Kampala, Uganda

Kennedy Muni[1*], Olive Kobusingye[2], Charlie Mock[1,3,4], James P Hughes[5], Philip M Hurvitz[6], Brandon Guthrie[1,3]

[1]Department of Epidemiology, University of Washington, Seattle, WA, USA

[2]Department of Disease Control and Environmental Health, Makerere University, Kampala, Uganda

[3]Department of Global Health, University of Washington, Seattle, WA, USA

[4]Department of Surgery, University of Washington, Seattle, WA, USA

[5]Department of Biostatistics, University of Washington, Seattle, WA, USA

[6]Department of Urban Design and Planning, University of Washington, Seattle, WA, USA

***Address correspondence to**: Kennedy Muni, Department of Epidemiology, 1959 NE Pacific Street, Health Sciences Bldg, F-262, Box 357236, Seattle, WA 98195. **Tel**.: 206-543-1065; **E-mail**: muni82@uw.edu

Key words: Helmet, training, motorcycle, cross-sectional study, driver

ABSTRACT

Background *SafeBoda* is a motorcycle taxi company that provides road safety training and helmets to its drivers in Kampala, Uganda. We sought to determine whether *SafeBoda* drivers are more likely to engage in safe riding behaviors than regular drivers (motorcycle taxi drivers not part of *SafeBoda*).

Methods We measured riding behaviors in *SafeBoda* and regular drivers through: a) computer-assisted personal interview (CAPI), where 400 drivers were asked about their riding behaviors (e.g., helmet and mobile phone use) and b) roadside observation, where riding behaviors were observed in 3000 boda-boda drivers and their passengers along major roads in Kampala.

Results Across the two cross-sectional studies, a higher proportion of *SafeBoda* drivers than regular drivers engaged in safe riding behaviors. For instance, helmet use among *SafeBoda* compared to regular drivers was 21 percent points higher (95% CI: 0.15-0.27; $p<0.001$) based on the CAPI and 45 percent points higher (95% CI: 0.43-0.47; $p<0.001$) based on roadside observation. Furthermore, compared to regular drivers, *SafeBoda* drivers were more likely to report having a driver's license (66.3% vs 33.5%; $p<0.001$) and a reflective jacket (99.5% vs 50.5%; $p<0.001$) and were less likely to report driving towards oncoming traffic (4% vs 45.7%; $p<0.001$) in the past 30 days.

Conclusion The *SafeBoda* program is associated with increased safe riding behaviors among motorcycle taxi drivers in Kampala. Therefore, the promotion and expansion of such programs may lead to a reduction in morbidity and mortality due to road injuries.

INTRODUCTION

In Uganda, motorcycle taxis known locally as boda-boda,[a] are a major form of transportation because they are relatively affordable, convenient for short distances, and efficient for navigating heavy vehicular traffic.[5,7,26-28] In the city of Kampala alone, there are an estimated 50,000 to 80,000 boda-boda drivers.[27] Despite their advantages, motorcycle taxis are involved in a substantial number of road traffic crashes (RTCs) and associated injuries in Uganda.[4,10,29,37,38] Boda-boda drivers are the second largest road user category involved in RTCs in Uganda. In 2015, they were involved in 28.2% of RTCs and represented 23.5% of road traffic injuries in Uganda.[10] According to the Road Safety Performance Review report, road fatalities involving boda-boda drivers doubled between 2011 and 2015 (from 570 deaths to 1170 deaths).[29] Moreover, Mulago, the main national public hospital in Uganda, receives between 10 and 20 cases of boda-boda related RTCs daily, and spends more than 60% of its annual surgery budget on treating trauma from these crashes.[27,38,39] Poor riding behaviors such as speeding, unsafe overtaking, and low use of protective gear have been found to be associated with increased crash and injury risk in this road user group.[7,26,29,37,38] One study conducted among boda-boda drivers presenting at the emergency unit at Mulago Hospital in Kampala, found 57% lower odds of head injury in drivers wearing a helmet during the crash compared to those not wearing a helmet.[38] Yet, despite this evidence, helmet use among boda-boda drivers and their passengers in Kampala is low. For instance, a 2011 roadside observation study of 12189 boda-boda drivers and passengers found helmet use of 30.8% and <1% respectively.[26]

In recent years, private companies such as *SafeBoda* and *Taxify* have cropped up to formalize the boda-boda sector and to address the issues around poor riding behaviors through training and provision of protective equipment.[31,40] *SafeBoda* is a road safety-conscious private

---

[a] This term originally referred to bicycle taxis that operated at the Uganda-Kenya border. However, it is now a generic term for bicycle and motorcycle taxis in East Africa. In present use, boda-boda increasingly refers to motorcycle taxis.

company founded in 2015. It currently provides motorcycle taxi services in Kampala using a trained community of drivers. The company provides multiphase road safety training, two helmets (for driver and passenger), and vehicle maintenance to its drivers in an effort to increase safe riding behaviors and reduce crashes and injuries. The company also provides hairnets to encourage helmet use in passengers concerned about contracting skin diseases from a shared helmet.[40] *SafeBoda* currently has over 5000 drivers in Kampala.

There is contradictory evidence on whether programs such as *SafeBoda* that provide protective gear and training to motorcycle drivers in low-and middle-income countries are associated with increased safe riding behaviors and reduced incidents of RTC and injuries.[2,6,34-36] Therefore, we sought to determine whether drivers enrolled in the *SafeBoda* program were more likely to engage in safe riding behaviors than regular drivers in Kampala, Uganda. In this study and throughout this paper, regular drivers refer to boda-boda drivers who were not identified as part of the *SafeBoda* program. At the time this study commenced, *SafeBoda* was the only company of its kind operating in Kampala. Since that time, additional companies (i.e., *UberBoda*, *Taxify*, and *Dial Jack*) have started to provide similar services in the city.

METHODS

We conducted two cross-sectional studies to measure safe riding behaviors (e.g., helmet use) in *SafeBoda* and regular drivers in Kampala.

Computer-assisted personal interview (CAPI) study: We recruited 400 drivers (200 *SafeBoda* and 200 regular drivers) between October and December 2017 from boda-boda stages across Kampala. A stage is a location where a boda-boda driver is generally stationed when not out driving a passenger or looking for one.

Study size: We used helmet use as the primary outcome to calculate the study sample size. Previous studies had reported helmet use between 18% and 30% among boda-boda drivers in

Uganda.[26,38] We assumed that 30% of regular drivers wore a helmet in Kampala. With a type I error rate of 0.05 using a χ^2 test, we needed a sample of approximately 400 boda-boda drivers to have 80% power to detect a prevalence ratio (PR) of 1.45 in helmet use comparing *SafeBoda* and regular drivers.[41] We used a purposive and convenience sampling method (i.e., targeting areas with boda-boda stages with at least one *SafeBoda* driver) to recruit participants for the study. This method had previously been used to recruit boda-boda drivers in Kampala.[5] We recruited a maximum of 4 drivers from each stage by enrolling the first 4 to give consent to be in the study. We limited the number recruited from each stage to avoid overrepresentation of any one stage in the study. Drivers were eligible to participate if they were 18 years or older, able to communicate in English or Luganda (a widely spoken language in Uganda), and had been working as a boda-boda driver for at least six consecutive months at the time of recruitment. The choice of six months was to help us to collect covariate information such as past crash history.

Data collection: Consenting drivers were administered a 15-minute questionnaire (Figure S1) in a reasonably quiet location near the recruitment stage. The questionnaire contained questions on demographic and personal characteristics (e.g., age, education, income, years of boda-boda experience, and hours worked per day) and on riding behaviors (e.g., helmet use, possession of a driver's license and reflective vest, and alcohol and mobile phone use). A complete list of the variables in the questionnaire can be found in the supplementary materials (Figure S1).

The exposure (*SafeBoda* status) was ascertained through self-report and/or by observing presence or absence of a *SafeBoda* branded jacket at the time of the interview. The primary outcome (frequency of helmet use in the past 30 days) was measured through self-report and was categorized in the analysis as "yes" for always, "no" for other responses). All variables were measured through self-report or direct observation.

The CAPI questions were derived from previous studies and discussions within the study team.[5,6,13-16,21] The questionnaire was available in English and Luganda.

Spotter observation study: In order to validate the self-reported helmet use from the CAPI study, we observed actual helmet use on leaving the boda-boda stage in 20% of the drivers in the CAPI study. Every 2nd CAPI study participant to be interviewed at each stage was systematically selected and observed for helmet use as they left the stage following the interview. Recruitment was continued until the sample size of 80 drivers was reached. Participants did not know they would be observed.

Roadside observation study: We collected data on *SafeBoda* status and helmet use and other riding behaviors (e.g., mobile phone use while driving and passenger helmet use) from 3000 boda-boda drivers along purposively selected streets in Kampala. Streets were selected based on traffic volume and to be representative of the five administrative divisions of Kampala (i.e., Rubaga, Kawempe, Kampala Central, Nakawa, and Makindye). *SafeBoda* status was straightforward to ascertain as *SafeBoda* drivers wear brightly colored orange reflective jackets inscribed with the company's logo. Regular boda-boda drivers were distinguished from other motorcyclists (i.e., non-boda-boda motorcyclists) by the smaller size of their motorcycle (engine size of 100-125 cc) and their demeanor (e.g., looking around for customers). We collected data on the first boda-boda driver to pass through the lane nearest to the study's observation point. This was done at two-minute intervals. If no boda-boda driver drove past the observation point during the first 30 seconds of the two-minute interval, no observation was made until the next interval. To capture both peak and non-peak hours, we collected data at each observation point for two peak hours (between 7 am and 9 am or 4 pm and 7 pm) and two non-peak hours (between 10 am and 3 pm) in one day. To minimize the risk of double counting, each selected street was observed in one travel direction, which was chosen by the flip of a coin.

Analysis: All analyses were conducted in the statistical package R (version 3.5.1).[42] We used Pearson's chi-square with continuity correction (for categorical variables) and Student's t-tests (for continuous variables) to compare outcomes (e.g., helmet ownership and past crash history) between *SafeBoda* and regular drivers. We then compared the proportion of helmet use by *SafeBoda* and regular drivers from the two studies. We used the *EpiStats* package in R to assess differences in proportions of self-reported helmet use, observed helmet use, and road traffic crashes between *SafeBoda* and regular drivers.[43] Other riding behaviors (e.g., alcohol use before driving and driving on pedestrian walkways) were analyzed as proportions.

Lastly, we ran a Poisson regression model to estimate the PR to assess whether being in the *SafeBod*a program was associated with increased self-reported helmet use while adjusting for a minimum set of potential confounders determined *a priori*. The minimum set of confounders (age, education, and driver's license) came from a directed acyclic graph of hypothesized variable relationships (informed by theory and evidence from existing injury literature). We opted to use a Poisson model instead of a logistic model due to helmet use not being a rare outcome in the boda-boda population. We used robust standard errors in the Poisson model to account for any violation of the distribution assumption that the variance equals the mean.

Sensitivity analysis: To examine the adequacy of the covariate adjustment model, we used another approach (propensity score matching) to control for confounding of the *SafeBoda*-helmet use association. The matched exposure sets were created using the *MatchIt* package in R.[44] Due to limitations in the *MatchIt* package, this analysis used only observations with complete data. We included in the matching, the exposure and variables believed to be associated with the exposure and outcome (i.e., age, education, and driver's license). A Poisson regression model adjusting for these confounders was then used to compare matched pairs of *SafeBoda* and regular drivers.

We conducted a second sensitivity analysis to assess how robust our self-reported helmet use data were against potential outcome misclassification due to social desirability bias. We imputed observed helmet use for drivers with self-reported helmet use but who were not observed in the spotter observation study (internal validation group). We took advantage of the relationships between observed helmet use, self-reported helmet use, *SafeBoda* status, and relevant covariates (e.g., age) in the validation group to impute the missing observed helmet use for drivers not in the validation group.[45,46] The *mice* package in R was used for the multiple imputation process.[47] We used a logistic regression model for monotone missing data to impute the missing observed helmet use.[45,48] Upon completion of the imputation process, we generated a single PR comparing imputed observed helmet use in *SafeBoda* and regular drivers by pooling the PRs from 10 imputation rounds (using a Poisson model). We compared this PR to the PR from the naïve (self-reported) and spotter observation analyses. All models adjusted for the same set of covariates.

Ethical approval and consent: Approval for the study was granted by University of Washington, Makerere University, and the Uganda National Council for Science and Technology. Participants in the CAPI study provided verbal consent to be interviewed and were compensated for their time.

RESULTS

Among 200 *SafeBoda* and 200 regular drivers who completed the CAPI, compared to regular drivers, *SafeBoda* drivers had more boda-boda experience, were more educated, and made more money per week (Table 1).

Table 1 Comparison of baseline characteristics of *SafeBoda* and regular boda-boda drivers in Kampala (CAPI study)

Characteristic	*SafeBoda* drivers N = 200	Regular drivers N = 200	P value
	n (%) or mean (SD)	*n (%) or mean (SD)*	
Age (years)	33.5 (7.2)	32.1 (7.0)	0.05
Education (years)	8.4 (3.7)	8.0 (3.6)	0.25
Weekly net income (in USD)	25.05 (14.19)	23.16 (12.33)	0.16
Number of trips per day	18.3 (7.9)	18.6 (7.7)	0.65
Hours worked as boda-boda per day	12.6 (2.2)	12.2 (2.0)	0.03
Has a driving license	132 (66.3%)	67 (33.5%)	<0.001
Has other job	64 (32%)	76 (38.0%)	0.25
Used a phone while driving in past 30 days	39 (19.5%)	92 (46%)	<0.001
Alcohol use in past 2 hours before driving in past 30 days	4 (12.1%)	26 (41.3%)	0.007
Had a road traffic crash in past 6 months	43 (21.5%)	67 (33.5%)	0.01
Own a helmet	198 (99%)	185 (92.5%)	0.003
Helmet cost (in USD)	27.47 (9.95)	10.02 (7.06)	<0.001
Has a reflective jacket	199 (99.5%)	101 (50.5%)	<0.001
Has ever received a road safety training	198 (100%)	146 (74.1%)	<0.001
Has driven on a pedestrian sidewalk in past 30 days	14 (7.0%)	91 (45.5%)	<0.001
Has driven towards oncoming traffic in past 30 days	8 (4.0%)	91 (45.7%)	<0.001
Carried more than one passenger in past 30 days	18 (9.1%)	156 (78.4%)	<0.001
Years working as a boda-boda driver	7.0 (4.2)	6.3 (4.1)	0.10

Note: Continuous variables are reported as mean with standard deviation (in parenthesis), while categorical variables are reported as count with percent (in parenthesis). Some observations have missing data.

In the roadside observation study of 3000 boda-boda drivers in Kampala, 49 (1.6%) were identified as *SafeBoda* drivers and 2951 (98.4%) were classified as regular drivers (Table 2). Only 57% of the drivers were observed to carry a passenger, of which 9% carried more than one passenger (none among *SafeBoda* and 9.5% among regular drivers). About two-thirds of the passengers carried were male and this was similar in both *SafeBoda* and regular drivers. A small percentage of drivers (1.2%) were observed to engage in distracted driving behavior by using their phone while driving (6.1% among *SafeBoda* vs 1.2% among regular drivers).

Table 2 Characteristics of boda-boda drivers in Kampala (roadside observation study)

Characteristic	*SafeBoda* drivers N = 49	Regular drivers N = 2951
	n (%)	*n (%)*
Division of Kampala		
Kampala Central	23 (46.9%)	794 (26.9%)
Nakawa	12 (24.5%)	805 (27.3%)
Makindye	10 (20.4%)	569 (19.3%)
Kawempe	3 (6.1%)	313 (10.6%)
Rubaga	1 (2.0%)	470 (15.9%)
Carried a passenger during observation	35 (71.4%)	1689 (57.2%)
Carried more than one passenger	0 (0.0%)	161 (9.5%)
Driver observed to be using their mobile phone	3 (6.1%)	34 (1.2%)
Sex of first boda-boda passenger		
Female	13 (37.1%)	622 (36.8%)
Male	22 (62.9%)	1067 (63.2%)
First passenger wore a helmet	8 (22.9%)	11 (0.7%)
First passenger helmet was correctly strapped	6 (75%)	8 (72.7%)

Note: a) Categorical variables are reported as count with percent (in parenthesis). Data exclude missing observations. b) Adjusting for division did not significantly change observed difference in helmet use in SafeBoda and regular drivers (PR: 1.81 vs 1.76)

Helmet use: We compared the proportion of helmet use in boda-boda drivers, as measured by self-report and direct observation. Of the 400 drivers who reported their frequency of helmet use in the 30 days before the study, 88% (95% CI: 84.3%-90.9%) reported always using a helmet while driving, compared to 69.6% (95% CI: 58.1%-79.2%) observed to use a helmet during the spotter observation of 80 participants in the CAPI. In the roadside observation study of 3000 boda-boda drivers, we observed helmet use by 56.0% (95% CI: 54.2%-57.8%) of drivers. Observed helmet use among passengers was very low at 1.1% (95% CI: 0.69%-1.8%).

Across the two studies, a higher proportion of *SafeBoda* drivers used a helmet while driving than did regular drivers (Table 3). The difference in the proportion of helmet use between *SafeBoda* and regular drivers was 0.21 (95% CI: 0.15-0.27; p<0.001) in the CAPI study and 0.45 (95% CI: 0.43-0.47; p<0.001) in the roadside observation study. From the roadside observation, the difference in the proportion of helmet use between passengers carried by

SafeBoda drivers and passengers carried by regular drivers was 0.22 (95% CI: 0.08-0.36; $p<0.001$).

Table 3 Proportion of helmet use for *SafeBoda* and regular drivers and passengers: results from two approaches

Approach	*SafeBoda*				Regular			
	Driver		**Passenger**		**Driver**		**Passenger**	
	n (N)	Proportion (95% CIs)	n (N)	Proportion (95% CIs)	n (N)	Proportion (95% CIs)	n (N)	Proportion (95% CIs)
Self-reported (CAPI)	197 (200)	0.99 (0.95, 1.00)	-	-	155 (200)	0.78 (0.71, 0.83)	-	-
Spotter observation (CAPI)	24 (25)	0.96 (0.78, 1.00)	-	-	31 (54)	0.57 (0.43, 0.70)	-	-
Roadside observation	49 (49)	1.00 (0.91, 1.00)	8 (35)	0.23 (0.11, 0.41)	1630 (2951)	0.55 (0.53, 0.57)	11 (1689)	0.007 (0.003, 0.012)

Upon covariate adjustment for potential confounders (Figure S2), *SafeBoda* drivers were 1.22 times more likely to use a helmet while driving than regular drivers (95% CI: 1.14-1.31; p<0.001) in the CAPI study (Table 4). Adjustment for confounding with propensity score matching gave similar results (PR: 1.21, 95% CI: 1.11-1.32; p<0.001). There were 129 matched exposure pairs from a dataset of 394 participants (Figures S3-S5).

Table 4 Adjusted prevalence ratios (PRs) for self-reported helmet use comparing *SafeBoda* and regular drivers

	PR (covariate adjustment)	**PR (propensity score matching)**
Regular boda-boda driver	1.00	1.00
***SafeBoda* driver**	1.22 (1.14-1.31)	1.21 (1.11-1.32)

Note: Adjusted for age, education, and possession of a driver's license

Other riding behaviors: Compared to regular drivers, *SafeBoda* drivers were more likely to self-report other safe riding behaviors in the CAPI study. They were more likely to report having a driver's license (66.3% vs 33.5%; p<0.001) and a reflective jacket (99.5% vs 50.5%; p<0.001). Moreover, they were less likely to report talking on the phone while driving (19.5% vs 46%; p<0.001), driving towards oncoming traffic (4% vs 45.7%; p<0.001), and carrying more than one passenger (9.1% vs 78.4%; p<0.001) in the past 30 days (Table 1). From the roadside observation, *SafeBoda* drivers were more likely to be carrying a passenger (71% vs 57%; p = 0.07), less likely to be carrying more than 1 passenger (0% vs 9.5%; p = 0.10), and more likely to be observed using a mobile phone while driving (6.1% vs 1.2%; p = 0.01) (Table 2).

Road traffic crashes: Compared to regular drivers, *SafeBoda* drivers were less likely to report having been involved in a road traffic crash in the six months prior to the CAPI study (RR: 0.64, 95% CI: 0.46-0.89; p = 0.007). The majority of the 110 reported crashes involved either a collision with another motorcycle (27.3%) or a car (53.6%). Speeding (38.2%), making an illegal turn (30.9%), and driving towards oncoming traffic (17.3%) were the most frequently mentioned causes of the reported crashes. Among those who reported a crash, 89 (80.9%) were injured[2] and 81% of these injuries required a visit to a health facility.

Sensitivity analysis: Compared to the effect size based on self-reported helmet use (PR = 1.22), we found higher PRs comparing *SafeBoda* and regular drivers based on observed helmet use. Using the spotter observation data, *SafeBoda* drivers were 61% more likely to use a helmet while driving than regular drivers (PR: 1.61, 95% CI: 1.26-2.04). Results from the multiple imputation were similar (PR: 1.60, 95% CI: 1.53-1.67) but with narrower confidence intervals.

DISCUSSION

Our findings suggest that drivers in the *SafeBoda* program were not only more likely to wear a helmet while driving but they were also more likely to engage in other safe riding behaviors such as wearing a reflective vest, and less likely to engage in risky behaviors such as driving towards oncoming traffic. However, observed mobile phone use while driving was higher in *SafeBoda* compared to regular drivers, although the observed number was quite small.

One mechanism that may explain the observed higher rate of safe driving behaviors and helmet use among *SafeBoda* drivers, is the possibility that the *SafeBoda* program may be recruiting more safety conscious drivers. However, other plausible mechanisms for the observed findings include the program's provision of road safety training and protective equipment to its drivers and the sense of community among *SafeBoda* drivers. Not only are the *SafeBoda* drivers provided with two helmets and a reflective vest, they are also trained to view themselves as a community and to look out for each other as they drive. They seem more likely to follow traffic regulations than regular boda-boda drivers because of their training but also because they know other *SafeBoda* drivers are watching them. The observed higher helmet use in *SafeBoda*

[2] Injury was defined as any wound or bruise from a road traffic crash regardless of whether treatment was sought.

passengers may be because helmets are provided by the program, whereas regular boda-boda passengers would need to purchase their own. In addition, messages urging helmet use appear on passenger phone screen when ordering for a *SafeBoda* via their app. However, the higher helmet use could also be because *SafeBoda* passengers are more safety conscious.

Our sensitivity analyses seem to support the possibility of outcome misclassification from social desirability bias. We found a higher PR for observed helmet use in the sensitivity analyses than in the primary analysis. However, a model with an interaction term for the exposure (*SafeBoda*) and misclassified outcome (self-reported helmet use) was not statistically significant. In addition, the similar PRs for the spotter observation and multiply imputed data likely point to missingness at random in our data. This discrepancy in self-reported and observed helmet use has been previously reported in other settings.[16,49,50]

Our findings corroborate those of previous investigators that found generally low helmet use and frequent high-risk driving behaviors among regular boda-boda drivers in Kampala.[5,26,38] Similar to a 2011 roadside observation study by Roehler and colleagues in Kampala and a 2017 study by Bachani et al. in Kenya, we found low helmet use (1.1%) in boda-boda passengers.[16,26] However, unlike Roehler et al., we found a much higher observed helmet use in boda-boda drivers in our roadside observation study (56% versus 31%).[26] This discrepancy could be due to differences in our sampling approaches or to an increasing trend of helmet use over time (especially with the advent of programs such as *SafeBoda*).

This study has some limitations. First, we relied on self-report for the majority of the variables in the analysis. This may have introduced bias in our study estimates. We attempted to control for this potential bias, in both the design and analytic phases, by including direct observation studies and conducting sensitivity analyses. Second, the numbers of *SafeBoda* drivers and passengers observed during the roadside observation study were quite small. This has implications on the precision and generalizability of our findings. In addition, since we did not use probability sampling for the CAPI study, the generalizability of our findings to the whole boda-boda population in Kampala might be limited. However, it was not possible to use probability sampling due to the lack of a sampling frame for boda-boda stages in Kampala and the study's focus on stages with at least one *SafeBoda* driver.

Third, there is a possibility for selection bias if boda-boda drivers recruited from stages are systematically different from those riding on the streets. However, we found similar results (i.e., more road safety consciousness in *SafeBoda* drivers) in both the CAPI and roadside observation studies. Lastly, there was potential to misclassify non-boda-boda motorcyclists (i.e., courier and delivery drivers and drivers of personal motorcycles) as regular boda-boda drivers during the roadside observation study. This could lead to an attenuated association given that non-boda-boda motorcyclists generally have higher rates of helmet use than regular drivers. However, we believe there was little potential for misclassification of non-boda-boda motorcyclists as regular boda-boda drivers because regular boda-boda drivers in Kampala have a unique profile, including the types of the motorcycle they use and their behavior on the road. Another related misclassification limitation was the difficulty in identifying whether or not a driver's or a passenger's helmet was correctly strapped during the roadside observation study. We attempted to mitigate this limitation by using two observers during the study.

CONCLUSION

Our findings suggest that the *SafeBoda* program is associated with increased usage of helmets for both drivers and passengers and also with other safe behaviors such as not riding with more than one passenger. Given these findings, we believe *SafeBoda* and similar programs that recently joined the boda-boda industry (e.g., *Taxify* and *UberBoda*) may be worth scaling up in order to accelerate adoption of safe riding behaviors among boda-boda drivers.

What's already known on the subject
• Helmets reduce risk of head injury during a crash. • Helmet use is low among boda-boda drivers in Uganda.

What this study adds
• Helmet use estimates after introduction of the *SafeBoda* program. • Helmet use estimates from two measurement approaches. • Safe riding behaviors disaggregated by type of boda-boda driver.

Acknowledgments: The authors are grateful to Grace Magambo, Joseph Mugisha, and Simon Okurut for their support throughout the study.

Contributors: All authors conceived the question and study design. KM and OK were involved with data collection. KM, BG, and JPH conducted statistical analysis. All authors were involved with writing the manuscript. All authors read and approved the final version of the manuscript for publication.

Funding: This work was supported by the Fogarty International Center of the National Institutes of Health under Grant [#R25 TW009345]. However, the content of this publication is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Competing interests: None. *SafeBoda* was informed about the study but the company had no influence on the study.

Ethics approval: The study was approved by University of Washington IRB, Makerere University School of Public Health IRB, and Uganda National Council for Science and Technology.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement: The data that support the findings of this study are available from the corresponding author, [KM], upon reasonable request.

Chapter 3: Motorcycle taxi program is associated with reduced risk of road traffic crash among motorcycle taxi drivers in Kampala, Uganda

Kennedy Muni[1*], Olive Kobusingye[2], Charlie Mock[1,3,4], James P Hughes[5], Philip M Hurvitz[6], Brandon Guthrie[1,3]

[1]Department of Epidemiology, University of Washington, Seattle, WA, USA

[2]Department of Disease Control and Environmental Health, Makerere University, Kampala, Uganda

[3]Department of Global Health, University of Washington, Seattle, WA, USA

[4]Department of Surgery, University of Washington, Seattle, WA, USA

[5]Department of Biostatistics, University of Washington, Seattle, WA, USA

[6]Department of Urban Design and Planning, University of Washington, Seattle, WA, USA

***Address correspondence to**: Kennedy Muni, Department of Epidemiology, 1959 NE Pacific Street, Health Sciences Bldg, F-262, Box 357236, Seattle, WA 98195. **Tel**.: 206-543-1065; **E-mail**: muni82@uw.edu

ABSTRACT

SafeBoda is a transportation company that provides road safety training and helmets to its motorcycle taxi drivers in Kampala. We sought to determine whether risk of road traffic crash (RTC) was lower in *SafeBoda* compared to regular (non-SafeBoda) motorcycle taxi drivers during a 6-month follow-up period. We collected participant demographic and behavioral data at baseline using computer-assisted personal interview, and occurrence of RTC every two months using text messaging and telephone interview from a cohort of 342 drivers. There were 85 crashes (31 in *SafeBoda* and 54 in regular drivers) during follow-up. Over the six-month follow-up period, *SafeBoda* drivers were 39% less likely to be involved in a RTC than regular drivers after adjusting for age, possession of a driver's license, and education (RR: 0.61, 95% CI: 0.39-0.97, p=0.04). These findings suggest that the *SafeBoda* program results in safer driving and fewer RTCs among motorcycle taxi drivers in Kampala.

Key words: crash, injury, boda-boda, road safety

INTRODUCTION

Motorcycles are a popular mode of transportation in many low-and middle-income countries (LMICs) such as Uganda because they are relatively affordable, convenient for short distances and poor roads, and are efficient for navigating heavy traffic in many developing cities.[4-9] Despite these benefits, motorcycles are involved in a substantial number of road traffic crashes (RTCs)[3] and injuries in LMICs.[1-3,10] In the past two decades, many studies have reported increasing involvement of motorcycle drivers and their passengers in RTCs, as the number of motorcycles exploded in Africa.[3,6,9,12-14]

In Uganda, motorcycle taxis, known locally as boda-boda, are a major form of transportation.[26-28] In the city of Kampala alone, the number of boda-boda drivers is estimated to be between 50,000 and 80,000.[27] Boda-boda drivers are the second largest road user category involved in RTCs in Uganda. They were involved in 28.2% of RTCs and represented 23.5% of road traffic injuries in 2015.[10] According to the 2018 Road Safety Performance Review report, road fatalities involving boda-boda drivers doubled between 2011 and 2016.[29] Risky riding behaviors such as speeding, unsafe overtaking, driving under the influence of drugs, failure to wear protective gear (e.g., helmet and reflective jacket), driving towards oncoming traffic, and distracted driving (e.g., mobile phone use) are believed to account for the high crash and injury rates in this road user group.[5-8,16-23,29,37,38] Other factors include limited training for boda-boda drivers, poor road conditions, poor vehicle conditions (e.g., faulty brakes and bad tires) and weak enforcement of traffic regulations (e.g., on driver licensing) by the police.[5,6,8,16-23,29]

In recent years, road safety-conscious companies such as *SafeBoda* and *Taxify* have cropped up to formalize the boda-boda sector in Uganda and to address the issues around poor riding

[3] RTCs are defined here as incidents where a motorcycle taxi driver collided with another vehicle (e.g., car, motorcycle, and bicycle), pedestrian, animal, or stationary object (e.g., tree or pole). Injury is defined as any wound or bruise due to a road traffic crash involving a motorcycle taxi driver on a public road.

behaviors through training and provision of protective gear.[30,31] *SafeBoda* is a transportation company that takes an innovative approach to providing motorcycle taxi services using a trained community of drivers in an effort to increase safe riding behaviors and reduce crashes and injuries. The company provides multiphase road safety training, helmets, vehicle maintenance, and basic first responder training to its drivers. It also provides hairnets to passengers who are concerned about contracting skin diseases from a shared helmet.[30] Initially, the training for the drivers was provided by the Uganda police and the Uganda Red Cross Society. However, this has since been transitioned to a team of *SafeBoda* trainers (part of the *SafeBoda* academy) who were trained by the Global Road Safety Partnership and the Uganda police and Red Cross. Newly recruited drivers are trained on traffic signs and symbols, traffic regulations, the *SafeBoda* code of conduct, emergency response, customer care, and how to use the *SafeBoda* app.[30,32,33] *SafeBoda* drivers undergo regular refresher trainings. As of January 2019, *SafeBoda* has over 8000 drivers in Kampala.[33] At the time this study commenced, *SafeBoda* was the only company of its kind operating in Kampala. Since that time, additional companies (e.g., *UberBoda*, *Taxify*, and *Dial Jack*) have started to provide similar services in the city.

There is inconclusive evidence on whether road safety programs geared towards motorcycle drivers in LMICs are associated with increased safe riding behaviors and reduced incidents of RTC and injuries.[2,6,34-36] Moreover, no studies have been conducted in Uganda looking at the role of recent start-ups (e.g., *SafeBoda, Taxify, and UberBoda*) on road safety in the motorcycle taxi sector. Therefore, we sought to determine whether the risk of self-reported RTC is lower in *SafeBoda* drivers than in regular drivers. In addition, we wanted to determine whether road traffic knowledge mediates the hypothesized relationship between the *SafeBoda* program and risk of RTC. In this study and throughout this paper, regular drivers refer to boda-boda drivers who were not identified as part of the *SafeBoda* program.

MATERIALS AND METHODS

We recruited 342 drivers (171 *SafeBoda* and 171 regular drivers) between October 2017 and January 2018 from boda-boda stages across the five administrative divisions of Kampala. A stage is a location where a boda-boda driver is generally stationed when not out driving a passenger or looking for one. The majority of the drivers (n=330) included in this study were recruited from a separate cross-sectional study on safe riding behaviors conducted by our team in Kampala.[51]

Sample size, sampling methodology, and recruitment

We estimated that 15% of regular boda-boda drivers would have a RTC during the 6-month follow-up period. Since there were no estimates on this measure from previous studies in Uganda, this estimate was based on police reports. According to the 2015 police report, boda-boda drivers were involved in 28.2% of all RTCs in Uganda that year.[10] We assumed that over a 6-month period, the risk of involvement in RTC would be 15%, about half of the 28.2% reported for boda-boda drivers for 2015. Therefore, with a type I error rate of 0.05 using a χ^2 test, we needed a sample of approximately 342 drivers (accounting for 15% attrition) to have 80% power to detect a risk ratio of RTC of 0.35 comparing *SafeBoda* drivers to regular drivers.[52] We used a purposive and convenience sampling method (i.e., targeting areas with boda-boda stages with at least one *SafeBoda* driver) to recruit boda-boda drivers to the study. This method has previously been used to recruit boda-boda drivers in Kampala.[5] We recruited a maximum of 4 drivers from each stage by enrolling the first 4 who were willing to consent to be in the study. This limit was used to avoid overrepresentation of any one stage in the study. Drivers were eligible to be in the study if they were 18 years or older, able to communicate in English or Luganda (a widely spoken language in Uganda), and had been working as a boda-boda driver for at least 6 consecutive months at the time of recruitment. The choice of 6 months was to help us to accurately ascertain covariate information such as past crash history. In addition, consenting drivers must have had a functioning mobile phone. This was to facilitate the collection of RTC and injury data via text messaging and a telephone interview (for those who reported a crash and/or injury). We asked participants to provide a secondary contact number to facilitate easy follow-up in the event of the participant losing their phone, incapacitation, or death.

Data collection

We collected data using four tools: a) computer-assisted personal interview (CAPI) questionnaire, b) road traffic knowledge questionnaire, c) SMS/text-based questionnaire, and d) telephone interview questionnaire.

CAPI questionnaire

Eligible and consenting drivers were administered a 15-minute questionnaire (Figure S1) in a reasonably quiet location near the recruitment stage. The questionnaire was available in English and Luganda and contained questions on demographic and personal characteristics (e.g., age, education, income, years of boda-boda experience, and hours worked per day) and on riding behaviors (e.g., helmet use, possession of a reflective vest, and alcohol and mobile phone use). A complete list of the variables in the questionnaire can be found in the supplementary materials (Figure S1).

The exposure (*SafeBoda* status) was ascertained through self-report and/or by observing presence or absence of a *SafeBoda* branded jacket at the time of the interview. All variables were measured through self-report or direct observation. The CAPI questions were derived from previous studies and discussions within the study team. Previous studies in sub-Saharan Africa have successfully used self-report to collect similar data.[5,6,13-16,21]

Road traffic knowledge questionnaire

In addition to the CAPI, we administered a 10-item knowledge questionnaire of common traffic signs and laws (taken from the Uganda Highway Code) to all cohort participants (Figure S2). Each item in the knowledge questionnaire had an equal score (10 points) for a total possible score of 100.

SMS/text-based questionnaire

We collected data on occurrence of RTC and/or injury via a text message every two months. We used a reputable company with experience in bulk text messaging in Uganda to facilitate the data collection process.[53] Enrolled participants received a text message on their 60th, 120th, and 180th day of follow-up asking: "In the past 2 months, were you involved in a road traffic crash while driving your motorcycle? In the past 2 months did you sustain any injury from a road traffic crash while driving your motorcycle?" Participants were requested to respond either "yes" or "no" by text message. After 24 hours, participants with ambiguous responses and those who had not responded to the text message were followed up via a telephone call to their primary and secondary contact numbers. For participants reported as incapacitated or dead, data on any RTC or injury during the two-month interval were collected from the secondary contact or boda-boda peers from the participant's stage. Given the possibility of multiple occurrences of RTC in the two-month interval, data were only collected for the most recent occurrence of RTC (i.e., one crash per interval for participants who reported a crash in an interval).

Telephone interview questionnaire

For participants who reported a crash and/or an injury in the SMS/text-based questionnaire, we conducted a follow-up telephone interview to collect more detailed data on the reported RTC/injury (e.g., location of crash, severity of injury, and hospitalization). For non-responders or participants reported as incapacitated or dead, the same follow-up procedures outlined for the SMS/text-based questionnaire were followed.

Analysis

All analyses were conducted in the statistical package R (version 3.5.1).[42] We used Pearson's chi-square with continuity correction (for categorical variables) and Student's t-tests (for continuous variables) to compare baseline variables (e.g., helmet use, net weekly income, and age) between *SafeBoda* and regular drivers and to identify correlates of RTC during the follow-up period. For the correlates, RTC was defined as one crash per participant at any interval during the 6-month follow-up period.

We used generalized estimating equations (GEE) models to test whether the *SafeBoda* program was associated with lower risk of road traffic crash during the six months of follow-up. The adjusted model included a minimum set of potential confounders (age, education, and possession of a driver's license) selected from a causal diagram of hypothesized variable relationships. For the GEE models, we assumed exchangeable covariance structure (i.e., uniform correlations cross time), which we thought was reasonable for our data. Nonetheless, to minimize the risk of invalid inference from a misspecified covariance model, we fitted models using robust standard errors. The unadjusted and adjusted models were fitted using the *gee* package in R.[54] We conducted an additional GEE analysis by including a lagged RTC variable as a predictor to investigate whether having a crash in a previous 2-month interval was associated with crash risk in a subsequent interval.

We used the *mediation* package in R to estimate the total, direct, and indirect effects of the *SafeBoda* program-RTC relationship in the mediation analysis. The *mediation* package uses the counterfactual (potential outcomes) framework that allows the decomposition of the total effect into direct and indirect effects even in the presence of interaction and non-linearity.[55] We fit two regression models with cluster robust standard errors to estimate the various effects: a linear model for the effect of the *SafeBoda* program on knowledge score, adjusting for age, education, and possession of a driver's license; and a Poisson model for the effect of the *SafeBoda* program on RTCs, conditioning on knowledge score, age, education, and possession of a driver's license. The RTC variable was defined as the total number of crashes that occurred during the 6-months of follow-up. The *mediation* package allowed us to use a Markov Chain Monte Carlo approach to conduct 1000 simulations in order to estimate the average direct and indirect effects by averaging over each possible exposure and confounder values.

Since we collected exposure/*SafeBoda* program data only at baseline, we conducted an intention-to-treat (ITT)-like analysis, without regard for cross-overs that may have occurred during the follow-up period.

Sensitivity analyses

Our GEE analysis assumed missingness to be completely at random (MCAR). We ran a multiple imputation model using variables associated with RTC (e.g., age, education, history of road safety training, years of boda-boda experience, and possession of a driver's license). We used the *mice* package in R to impute missing data.[47] We imputed missing RTC data for participants lost to follow-up as well as any missing data in the other variables in the analytic model (i.e., age, education, and possession of a driver's license). The outcome (RTC) was modeled using logistic regression. All models used 10 imputation rounds and 30 iterations. We re-ran the GEE models using the imputed data. This was to gauge how robust our findings were to the MCAR assumption. We then compared the pooled risk ratio from the imputation to the estimate from the complete case analysis.

We conducted two additional sensitivity analyses: a) we assumed everyone that was lost to follow-up had a RTC in the intervals in which they were lost, b) we assumed everyone that was lost to follow-up did not have a RTC in the intervals in which they were lost.

Ethical approval and consent

Approval for the study was granted by University of Washington, Makerere University, and the Uganda National Council for Science and Technology. Participants provided verbal consent to be interviewed and were compensated for their time.

RESULTS

We enrolled a total of 342 boda-boda drivers in our cohort, among whom the mean age was 32.8 years, mean weekly net income was 23.5 US dollars, and mean years of working in the boda-boda sector was 6.7 years. In addition, the 342 boda-boda drivers had a mean of 8 years of schooling and the majority (86.4%) reported previously attending a road safety training. The drivers worked a mean of 12.4 hours and made 18.5 trips per day. Overall, 89.2% of the drivers reported always wearing a helmet while driving in the past 30 days although 96% said they owned a helmet which they had had for a mean of 15 months. The most frequently observed helmet type was full-face (68.5%). The majority of boda-boda drivers (73.3%) took less than 10 weeks from learning to drive to carrying a passenger. Sixty-three percent learned how to drive through a friend, relative or peer.

Among the 171 *SafeBoda* and 171 regular drivers who completed the CAPI questionnaire, compared to regular drivers, *SafeBoda* drivers had more years of boda-boda experience, were more educated, and made more money per week (Table 1). For instance, at enrollment compared to regular drivers, *SafeBoda* drivers were more likely to report safe riding behaviors (e.g., they were more likely to report having a driver's license (65.9% vs 34.5%, $p<0.001$) and wearing a reflective jacket (100% vs 49.1%, $p<0.001$). They were also less likely to report having been involved in a road traffic crash in the 6 months prior to their enrollment in the study (RR: 0.68, 95% CI: 0.48-0.97, p = 0.03). Detailed results comparing safe riding behaviors at enrollment and occurrence of road traffic crashes, injuries, and hospitalizations 6 months prior to enrollment in *SafeBoda* and regular drivers are reported elsewhere.[51]

Table 1 Comparison of baseline characteristics of *SafeBoda* and regular boda-boda drivers in Kampala

Characteristic	*SafeBoda* drivers N = 171	Regular drivers N = 171
Age (years)	33.3 (7.1)	32.4 (7.3)
Education (years)	8.2 (3.7)	8.0 (3.6)
Weekly net income (in USD)	24.4 (13.5)	22.6 (12.4)
Number of trips per day	18.0 (8.1)	19.0 (7.9)
Hours worked as boda-boda per day	12.6 (2.3)	12.1 (1.9)
Has a driving license	112 (65.9%)	59 (34.5%)
Has other job	60 (35.1%)	67 (39.2%)
Used a phone while driving in past 30 days	34 (19.9%)	84 (49.1%)
Alcohol use in past 2 hours before driving in past 30 days	3 (11.1%)	22 (41.5%)
Reported always using a helmet in past 30 days	169 (98.8%)	136 (79.5%)
Own a helmet	169 (98.8%)	158 (92.4%)
Helmet cost (in USD)	27.6 (10.5)	10.1 (7.5)
Has a reflective jacket	171 (100%)	84 (49.1%)
Has ever received a road safety training	169 (100%)	124 (72.9%)
Has driven on a pedestrian sidewalk in past 30 days	12 (7.0%)	74 (43.3%)
Has driven towards oncoming traffic in past 30 days	7 (4.1%)	78 (45.6%)
Carried more than one passenger in past 30 days	16 (9.4%)	134 (78.4%)
Years working as a boda-boda driver	7.1 (4.4)	6.4 (4.3)

Note: Continuous variables are reported as mean with standard deviation (in parenthesis), while dichotomous variables are reported as count with percent (in parenthesis). Some observations have missing data.

Risk of road traffic crash during the 6-month follow-up period

At the first follow-up, 337 (98.5%) drivers provided data on occurrence of RTC during the initial 2 months of follow-up. At 4 months, follow-up was successful for 334 (97.7%) drivers and at 6-months, 330 (96.5%) drivers completed follow-up. Over the course of the study, 12 (3.5%) drivers were lost to follow-up. There were 24 RTCs during the first 2 months of follow-up, 32 RTCs during the next 2 months of follow-up, and 29 RTCs during the last 2 months of follow-up. Overall, 69 unique RTCs (i.e., only one crash per driver for those who reported a crash during the follow-up) occurred over the entire follow-up period. Two drivers reported having 3 RTCs during the follow-up period (i.e., one RTC during each 2-month interval) and 14 drivers reported having at least 2 RTCs during the follow-up period (Figure S3).

Based on the unadjusted GEE model, there was a 43% lower risk of RTC during the 6-months of follow-up among *SafeBoda* compared to regular drivers (RR: 0.57, 95% CI: 0.36-0.89, p = 0.02) (Table 2). After adjusting for potential confounders (i.e., age, possession of a driver's license, and education), *SafeBoda* drivers were 39% less likely to be involved in a road traffic crash than regular drivers during the 6 months of follow-up (RR: 0.61, 95% CI: 0.39-0.97, p = 0.04).

Table 2 Crude and adjusted RTC risk associated with the *SafeBoda* program (results from the GEE & mediation models)

		Unadjusted model (95% CI)	Adjusted model (95% CI)*	Imputed unadjusted model (95% CI)	Imputed adjusted model (95% CI)*	Adjusted mediation model (95% CI)*
		Risk ratios (RR)				**Risk difference (RD)**
Type of boda-boda driver	**Regular**	Ref	Ref	Ref	Ref	Ref
	SafeBoda	0.57 (0.36-0.89)	0.61 (0.39-0.97)	0.55 (0.35-0.87)	0.62 (0.39-0.98)	-0.04 (-0.08, -0.01)

* adjusted for age, education, and driver's license

Types of crashes, severity of injuries, and hospitalization during the 6-month follow-up period

The majority of the 85 crashes (31 in *SafeBoda* vs 54 in regular drivers) reported during the follow-up period involved either a collision with another motorcycle (27.1%) or a car (63.5%). Speeding (10.6%), faulty brakes (7.1%), distracted driving (4.7%), and slippery surface (4.7%) were the most frequently mentioned causes of the reported crashes. For the 85 crashes, 69 (81.2%) resulted in injury to the driver and 56 (81.2%) of these driver injuries required a visit to a health facility. Of the injuries that required a visit to a health facility, 13 (23.2%) required in-patient care (admission). The median hospitalization time for injuries requiring in-patient care (at the time of data collection) was 3 days with a range from 1 day to 30 days.

Correlates of road traffic crash during the 6-month follow-up period

Occurrence of road traffic crash during the follow-up period was statistically significantly associated with age, ever use of alcohol, driving into oncoming traffic in the past 30 days, number of years working as a boda-boda driver, having had a crash in the 6 months prior to enrollment in the study, and ownership of a reflective vest. However, having had a crash in a previous 2-month interval was not statistically significantly associated with a crash risk in a subsequent interval (RR: 1.17, 95% CI: 0.58-2.35, p = 0.66), although the study was underpowered to detect such an association.

Mediation analysis

The mean road traffic knowledge score for all drivers was 86.3 (87.7 for *SafeBoda vs* 84.8 for regular drivers), with a statistically significant mean difference of 2.9 points (95% CI: 0.5-5.3, p=0.02). However, we did not find evidence of a statistically significant mediation effect of road traffic knowledge on risk of road traffic crash (Table 3). The effect of the *SafeBoda* program over the 6 months of follow-up through pathways involving road traffic knowledge was an increase of 1.36 crashes per 1,000 drivers (95% CI: -0.002 to 0.010) in *SafeBoda* drivers compared to regular drivers (p = 0.53), while the effect through pathways other than road traffic knowledge was a reduction of 43.0 crashes per 1,000 drivers (95% CI: -0.084 to -0.010, p = 0.03). The total effect of the *SafeBoda* program on risk of RTC (irrespective of pathway) was a reduction of 41.6 crashes per 1,000 drivers compared to regular drivers (95% CI: -0.081 to -0.010s, p = 0.03).

Table 3 Summary of results from the mediation analysis

	Mediation (RD scale, 95% CI)
Average direct effect, ζ	-0.0430 (-0.0840, -0.0100)
Average indirect effect, δ	0.0014 (-0.0023, 0.0100)
Total effect, τ	-0.0416 (-0.0806, - 0.0100)

Sensitivity analyses

Compared to the effect size based on the complete case analysis (adjusted RR: 0.61, 95% CI: 0.39-0.97, p = 0.04), we found a similar estimate for risk of RTC (adjusted RR: 0.62, 95% CI: 0.39-0.98, p = 0.04) from the multiply imputed data (Table 2). Models assuming all participants lost to follow-up either had a RTC or did not have a RTC in the intervals in which they were lost, found similar effect estimate (data not shown). Moreover, we found little missingness (Table 4) in our data (i.e., 1.5% for RTC in the initial 2 months of follow-up, 2.3% at 4 months, and 3.5% at 6 months).

Table 4 Reasons for dropout from the study among participants lost to follow-up

Reason for dropout from study	**Number (%)**
Left the country or study area	4 (33.3%)
Died from non-study-related causes	2 (16.7%)
Could not be reached/not found at the stage	6 (50%)
Total	12

DISCUSSION

Our findings suggest that drivers in the *SafeBoda* program are not only more likely to engage in safe riding behaviors (e.g., wearing a helmet and a reflective vest while driving and not carrying more than one passenger at a time), but also have fewer RTCs than regular drivers.

There are multiple mechanisms that may explain the observed higher frequency of safe driving behaviors and reduced risk of RTC among *SafeBoda* drivers. Plausible mechanisms include the self-selection of more safety conscious drivers into the *SafeBoda* program, the provision of road safety training, bike maintenance, and protective equipment by the program to its drivers, the program's code of conduct that includes penalties (e.g., re-training, suspension, and dismissal) for unsafe driving behaviors, and the sense of community among *SafeBoda* drivers. *SafeBoda* drivers undergo a rigorous road safety training and are trained to view themselves as a community of drivers that self-regulate as they drive. Anecdotal evidence indicates that they are more likely to stop at traffic stops than regular boda-boda drivers, not only because they are trained to obey traffic regulations but also because they know other *SafeBoda* drivers are watching them (there are penalties for drivers reported to have engaged in unsafe behaviors (e.g., running a traffic stop) and poor customer care. In this study, we investigated how one of these mechanisms (changes in road traffic knowledge due to road safety training) might impact occurrence of road traffic crashes but were unable to find evidence that the effect of the *SafeBoda* program is mediated through traffic safety knowledge. It is plausible this null result is due to the relatively high road traffic knowledge score for regular drivers (84.8). A high proportion (73%) of regular boda-boda drivers reported having received road safety training in the past (which might explain the high knowledge score), however, these trainings seemed to have had limited effect on their riding behaviors and risk of RTC.

Our findings corroborate those of previous investigators that found safer riding behaviors and reduced risk of RTCs and injuries among motorcycle drivers who have undergone road safety training.[5,6,34,56,57] For example, an evaluation of a road safety program (the Asia-Pacific Honda Safety Riding Program) in Thailand, reported 29% lower odds of road traffic injuries in the trained group of motorcycle drivers compared to the controls (who did not receive road safety training).[34] Similarly, an evaluation of the *Tukbodabike* program (a boda-boda-led training initiative in Western Kenya) found 6% lower incidence of RTCs among trained drivers compared to untrained ones.[6] Other studies have however found no evidence of reduced risk of RTCs and injuries upon undergoing road safety training.[2,35,36]

Our study has some limitations. First, we relied on self-report for the majority of variables in the analysis, including the main outcome (road traffic crash) and confounders (e.g. possession of a driver's license). This may have introduced measurement error into our study estimates, especially if social desirability bias in reporting occurrence of road traffic crash differed in *SafeBoda* and regular drivers. It is possible that *SafeBoda* drivers may be less likely to report a RTC compared to regular drivers. Moreover, reporting of RTCs may be biased towards more serious crashes (with less serious crashes being unreported), so our estimates are likely to be underestimates. The fact that 81.2% of those who had a crash were injured reinforces this possibility. However, reliance on self-reports was unavoidable because there are currently no comprehensive national crash and injury registries in Uganda.

Second, since we did not use probability sampling to recruit participants into the study, the generalizability of our findings to the whole boda-boda population in Kampala might be limited. However, it was not possible to use probability sampling due to the lack of a sampling frame for boda-boda stages in Kampala and the study's focus on stages with at least one *SafeBoda* driver.

Third, we conducted an intent-to-treat analysis without regard for cross-overs that may have occurred during the follow-up period. Given that the *SafeBoda* program expanded dramatically during the study period (i.e., from about 1,200 drivers at the time of recruitment to over 4,000 drivers at the time the follow-up period ended), it is plausible that some regular drivers in our cohort may have become *SafeBoda* drivers during the follow-up (however, we did not collect any data on cross-overs). This would bias our findings towards a null association especially if a large number of regular drivers crossed over to the *SafeBoda* program during the follow-up period.

Lastly, our mediation analysis relied on the potential outcomes framework and required that certain causal inference assumptions be met in order to identify the true causal effect. (e.g., well defined intervention, no unmeasured exposure-outcome confounders, no unmeasured mediator-outcome confounders, and no unmeasured exposure-mediator confounders). These assumptions may be too strong to apply to a multi-component intervention such as the *SafeBoda* program and a dataset from an observational study.

Despite these limitations, our findings suggest that the *SafeBoda* program is associated with increased safe riding behaviors and reduced risk of road traffic crash among boda-boda drivers. The reduced risk of RTCs was shown both in terms of past and future risk of RTC with similar magnitude (at least 30% reduction) and the association was robust to sensitivity analyses. Given these findings, we believe the model of the *SafeBoda* and other similar programs (e.g., *Taxify* and *UberBoda*) should be used to inform injury prevention policy (especially around training and licensing of boda-boda drivers). In addition, our findings show that these programs may be worth scaling up in order to accelerate adoption of safe riding behaviors in boda-boda drivers in Uganda and globally.

Acknowledgments

The authors are grateful to Grace Magambo, Joseph Mugisha, and Simon Okurut for their support throughout the study.

Declaration of interest statement

The authors declare no competing interests.

Funding

This work was supported by the Fogarty International Center of the National Institutes of Health under Grant [#R25 TW009345]. However, the content of this publication is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data availability statement

The data that support the findings of this study are available from the corresponding author, [KM], upon reasonable request.

Contributors

All authors conceived the question and study design. KM and OK were involved with data collection. KM, BG, and JPH conducted statistical analysis. All authors were involved with writing the document. All authors read and approved the final version of the manuscript for publication.

Chapter 4: Using Global Positioning System (GPS) devices to measure driving behaviors in motorcycle taxi drivers in Kampala: an exploratory naturalistic study

Kennedy Muni[1*], Philip M Hurvitz[2], Olive Kobusingye[3], Charlie Mock[1,4,5], James P Hughes[6], Brandon Guthrie[1,4]

[1]Department of Epidemiology, University of Washington, Seattle, WA, USA

[2]Department of Urban Design and Planning, University of Washington, Seattle, WA, USA

[3]Department of Disease Control and Environmental Health, Makerere University, Kampala, Uganda

[4]Department of Global Health, University of Washington, Seattle, WA, USA

[5]Department of Surgery, University of Washington, Seattle, WA, USA

[6]Department of Biostatistics, University of Washington, Seattle, WA, USA

***Address correspondence to**: Kennedy Muni, Department of Epidemiology, 1959 NE Pacific Street, Health Sciences Bldg., F-262, Box 357236, Seattle, WA 98195. **Tel**.: 206-543-1065; **E-mail**: muni82@uw.edu

Key words: GPS, motorcycle, cross-sectional study, driver

ABSTRACT

Background *SafeBoda* is a motorcycle taxi company that provides road safety training and helmets to its drivers in Kampala, Uganda. Evidence is needed about the effect of such programs on actual driving behaviors. We conducted an exploratory naturalistic study to assess the feasibility of using global positioning system (GPS) devices to measure movement patterns and driving behaviors (e.g., time spent riding, speed, total distance traveled) in a sample of motorcycle taxi drivers in Kampala. In addition, we sought to determine if these patterns and behaviors differ between *SafeBoda* and regular (not part of *SafeBoda*) drivers.

Methods We measured driving behaviors in 60 *SafeBoda* and regular drivers using a computer-assisted personal interview and a GPS device (where the driving patterns for each driver were recorded in 1 second intervals over a 24-hour period as the driver went about his daily travels). Trip summaries and heat maps were generated from point and segment data and GPS coordinates using PostGIS and the statistical package R.

Results We found that GPS devices can be used for measuring motorcycle taxi driver movements without any major measurement issues (e.g., missing data and device failure) and logistical problems (e.g., installation and retrieval of the devices). The GPS data showed that motorcycle taxi drivers in the study made an average of 31 trips per day (SD = 10.4). The median trip was 3.5 km and lasted on average 9.0 minutes (range 3.0 minutes to 13.8 minutes). The mean farthest distance (Euclidean) of a trip from the driver's stage or taxi stand was 5.8 km. The mean speed on a trip was 22.5 km/h. Measured driving behaviors did not differ noticeably between *SafeBoda* and regular drivers and the two groups of drivers appeared to share similar activity spaces within the city.

Conclusion Use of GPS devices for measuring driving patterns in motorcycle taxi drivers is feasible and provides another alternative to self-reports for measuring driver movements and driving behaviors.

INTRODUCTION

Motorcycles are a popular mode of transportation in many low-and middle-income countries (LMICs) such as Uganda because, compared to other vehicular transportation modes, they are relatively affordable, convenient for short distances and poor roads, and are efficient for navigating heavy traffic in many developing cities.[4-9] Despite these benefits, motorcycles are involved in a substantial number of road traffic crashes (RTCs) and injuries in LMICs.[1,3,10]

In Uganda, motorcycle taxis, known locally as boda-boda, are a major form of transportation.[26-28] In Kampala, the capital city of Uganda, the number of boda-boda drivers is estimated to be between 50,000 and 80,000, although this is likely an underestimate.[27] Many boda-boda drivers engage in risky driving behaviors (e.g., wrongful overtaking, driving on pedestrian walkways, and driving against traffic).[7,29,37,38] However, in recent years, road safety-conscious companies such as *SafeBoda, uberBoda,* and *Taxify* have cropped up to formalize the boda-boda sector in Uganda and to address issues related to poor driving behaviors through training and provision of protective gear.[30,31,58] *SafeBoda* is a transportation company that provides motorcycle taxi services using a trained community of drivers in an effort to increase safe driving behaviors and reduce crashes. The company provides multiphase road safety training, helmets, and vehicle maintenance to its drivers.[30] Initially, the training for drivers was provided by the Uganda Police and the Uganda Red Cross Society. However, this has since transitioned to a team of *SafeBoda* trainers (part of the *SafeBoda* academy) who were trained by the Global Road Safety Partnership and the Uganda Police and Red Cross. Newly recruited drivers are trained on traffic signage, traffic regulations, the *SafeBoda* code of conduct, emergency response, customer care, and how to use the *SafeBoda* app.[30,32,33] *SafeBoda* drivers undergo regular refresher trainings. As of January 2019, *SafeBoda* has over 8000 drivers in Kampala.[33] At the time this study commenced, *SafeBoda* was the only company of its kind operating in Kampala. Since that time, additional companies (e.g., *uberBoda*, *Taxify*, and *Dial Jack*) have started to provide similar services in the city.

Global positioning system (GPS) uses a constellation of satellites to determine the position of a ground receiver, which may be a standalone GPS device, a smartphone or any device with a GPS chipset.[59] Although GPS devices have been widely used to document movements patterns in both animals and humans,[59-63] their use to measure driving behaviors in motorcycle drivers has been limited, especially in low-and middle-income settings.[63-66] Historically, data on driving behaviors (which are important for understanding risk behaviors for RTCs and therefore informing interventions) in this road user group have been collected via self-reports. However, as GPS-enabled devices, particularly smartphones, become more readily affordable and available in low-resource settings, their use for collection of travel data, is becoming an interesting area of investigation. Moreover, use of GPS devices might help overcome measurement error issues associated with social desirability bias, non-validated tools, and literacy and language limitations inherent in self-report data in road safety research in low-resource settings.[61]

To the best of our knowledge, there has been only one published study that utilized GPS devices to measure driving behaviors among boda-boda drivers in Uganda.[65] However, similar studies have been conducted in other settings.[64,66] In their study in Uganda, Evans and colleagues used GPS devices and field interviews to understand boda-boda use in Kampala especially its role as a form of transportation and livelihood in the city.[65] However, the study had limited focus on road safety and did not investigate how driving patterns might differ by type of boda-boda driver, and provided very little detail on study logistics with respect to GPS measurement (e.g., device types, sampling interval, post-processing). Therefore, we sought to conduct an exploratory naturalistic study to assess the feasibility of using GPS devices to measure movement patterns and driving behaviors (e.g., time spent driving, speed, total distance traveled) in a sample of boda-boda drivers in Kampala. Additionally, we sought to determine if these patterns and behaviors differ between *SafeBoda* and regular boda-boda drivers. Naturalistic studies allow the unobtrusive use of a recording device to collect data on driver behavior and vehicular and environmental characteristics.[61] In this study and throughout this paper, regular drivers refer to boda-boda drivers who were not identified as part of the *SafeBoda* program.

METHODS

Sample size, sampling methodology, and recruitment

We recruited 60 drivers (30 *SafeBoda* and 30 regular drivers) between April 2018 and May 2018 from boda-boda stages across the five administrative divisions of Kampala (i.e., Kampala Central, Nakawa, Kawempe, Rubaga, and Makindye). A stage is a location where a boda-boda driver is generally stationed when not out driving a passenger or looking for one. All the 60 drivers included in this study came from a separate cross-sectional study on safe driving behaviors conducted by our team in Kampala.[51] The original study used a purposive and convenience sampling method (i.e., targeting areas with boda-boda stages with at least one *SafeBoda* driver) to recruit boda-boda drivers to the study. This method has previously been used to recruit boda-boda drivers in Kampala.[5]

For this study, we used stratified random sampling to recruit the 60 study participants. The boda-boda stages from the original study were stratified by the five administrative divisions of Kampala and then from each division, four stages were randomly selected for inclusion in the study. Once the stages were selected, all boda-boda drivers in each stage (between 1-4 boda-boda drivers) who had participated in the original study were asked if they would be interested in participating in the GPS study.

In addition to the eligibility criteria from the original study (i.e., participants must be 18 years or older, able to communicate in English or Luganda [a widely spoken language in Uganda], and had been working as a boda-boda driver for at least 6 consecutive months at the time of recruitment), participants consenting to the GPS study had to be willing to have the GPS device on them while driving their motorcycle for the study duration of 24 hours.

Data collection

We collected data using two instruments: a) a computer-assisted personal interview (CAPI) questionnaire, and b) a GPS device. Both instruments were pretested before actual data collection started.

CAPI questionnaire

Participants who consented to the GPS study were previously administered a 15-minute questionnaire (Figure S1) as part of a cross-sectional study on safe driving behaviors among boda-boda drivers in Kampala, Uganda.[51] The questionnaire was available in English and Luganda and contained questions on demographic and personal characteristics (e.g., age, education, income, years of boda-boda experience, and hours worked per day) and on driving behaviors (e.g., helmet use, possession of a reflective vest, and alcohol and mobile phone use). A complete list of the variables in the questionnaire can be found in the supplementary materials (Figure S1). The CAPI questions were derived from previous studies and discussions within the study team. Previous studies in sub-Saharan Africa have successfully used self-report to collect similar data.[5,6,13-16,21]

GPS devices

We used a commercial GPS device (QStarz BT-Q1000XT Bluetooth Data Logger GPS Receiver, Taipei, Taiwan) to record driver movement data in 1 second intervals as the driver went about their daily travels. The GPS devices were pre-programmed to record spatial and temporal variables (e.g., date, time, distance, latitude, longitude, and speed) as well as a variety of data quality variables (e.g. number of satellites (NSAT) and horizontal dilution of precision (HDOP)).[67-69] The movements of each participant were stored in the form of a trajectory which is defined as an ordered sequence of spatio-temporal positions (coordinates and timestamps).[67] Data were collected in Kampala over 7 weeks between April 2018 and May 2018 for a duration of about 24 hours per participant. Data collection was restricted to 24 hours because effective device memory was limited by the high frequency of collected data. Participants were instructed to put the GPS device in their jacket or trouser pocket and to drive normally for the entire duration of data collection. In addition, they were instructed to turn off the device when the 24 hours of recording elapsed. A research assistant called the participant 10-30 minutes before the end of the data collection period to remind him to turn off the device. Upon completion of data collection, the device was retrieved from the participant, and the GPS data downloaded (using the QTravel utility program that came with the devices) as comma-separated values (CSV) files, which were uploaded to a PostgreSQL/PostGIS database in preparation for processing and analysis.

Data processing

We conducted data processing[70] in PostgreSQL 11.2/PostGIS 2.5.1 and in the statistical package R.[71-73] Two different approaches (manual and algorithm-driven) were used to prepare the data for analysis, one using a manual, interactive approach using desktop GIS software, and the other using an automated approach using R scripts.

Manual approach: Two members of the study team jointly visually reviewed each participant's trajectory in QGIS to identify locations that appeared to be clustered points consistent with stops.[74] Clusters of points consistent with a stop were removed in order to select points that were part of trips (to facilitate calculation of trip characteristics such as mean trip speed). The team systematically identified clusters/data points for removal as follows: a) records for periods of inactivity (i.e., when GPS points appeared to be clustered around the boda-boda driver's home or stage), b) apparent clusters of records around a location indicating lack of movement/longer stops (e.g., stopped for a lunch break), and c) outliers or values that seem implausible (e.g., instantaneous speed greater 100 km/h).

Algorithm in R: We used a rule-based approach (algorithm) to stratify the continuous stream of GPS data into "stops" and "trips".[75] Stops were defined as relatively stationary clusters of points with low speed (i.e., <5 km/h). Any records that were not contained within a stop group were considered to be trips (regardless of whether there was a passenger involved). The algorithm processed the raw data as follows:

1. Flag records as being "slow" or "fast" (i.e., part of a stop or a trip) using a default cut-point of 5 km/h. This cut-off was chosen after discussions within the study team about what might constitute a reasonable lower speed threshold for boda-boda drivers on a trip.

2. Re-mark as "stop" those records that were greater than 5km/h but in a small set of temporally contiguous records (default threshold of 10 records) interspersed between groups of slow records. These records were likely the result of artifacts in the data due to errors in positioning or drivers moving relatively quickly for very short bursts.

3. Flag groups of temporally contiguous slow records (default threshold of 45 seconds) as being a "stop". The duration of 45 seconds was selected based on the study authors' knowledge of boda-boda movements in Kampala and visual inspection of the data.

4. Re-scan for "fast"/trip groups with duration less than 3 minutes (180 seconds) and recode as "stops". The study team made the decision to have 3 minutes as the minimum

threshold for trip time to accommodate stops at traffic lights. We conducted a sensitivity analysis that assessed how our decisions on the thresholds above might have impacted the trip summary findings.

Trip data summaries: We measured the distance (Euclidean) for each GPS point to the respective driver's stage using the PostGIS ST_Distance function with inputs being the GPS and stage point geometries (coordinates). Measurements were taken in meters using UTM zone 36 N (EPSG code 32636). We converted the point data to trip segments by creating a linestring between each temporally-sequential pair of records within a trip. For each segment, speed was calculated as distance/duration. Segments were aggregated by trip identifier for calculation of total duration and distance per trip. The R code used for data processing and analysis can be found in the supplementary materials (Figure S2).

Analysis

All analyses were conducted in PostgreSQL 11.2 (with PostGIS 2.5.1) and the statistical package R (version 3.5.2).[42,72,73] For the CAPI data, we used Pearson's chi-square with continuity correction (for categorical variables) and Student's t-tests (for continuous variables) to compare variables (e.g., time spent working as a boda-boda driver per day, helmet use and past crash history) between *SafeBoda* and regular boda-boda drivers.

We summarized trip data for each driver using a combination of point and segment data. For each trip the following summary statistics were calculated:

1. start timestamp
2. end timestamp
3. duration in seconds

4. minimum, maximum, mean, standard deviation of distance to stage

5. minimum, maximum, mean, standard deviation of segment speed

6. duration above 50 km/h

We then compared trip data by driver type (i.e., *SafeBoda* and regular drivers) using Student's t-test (for characteristics of the driver such as number of trips and total distance covered per day) and generalized estimating equations (GEE) models with a Gaussian link (for characteristics associated with trip such as trip duration and trip speed which are clustered within driver). We used the *gee* package in R for fitting the GEE models.[54]

Heat maps: Using the recorded GPS coordinates (longitude and latitude), we generated heat map rasters using the kde2d function in the R package *MASS*.[76] Heat maps were created separately for all GPS points as well as stratified by driver type. Data were resampled to 1-minute intervals to decrease processing time. The heat maps allowed visualization of which areas of Kampala were most frequented by the boda-boda drivers in the study.

Additionally, we created a heat map of speeding hotspots using data with speeds greater 50 km/h, the legal speed limit for urban and built areas in Uganda.

Activity space with kernel density estimator (KDE) analysis: The sp.kde function from the R package *spatialEco* was used to generate a kernel density estimator for each driver's GPS data, scaled from 0 to 1 with the "standardize = TRUE" option.[77] The cells within the 95% probability distribution were converted to polygons for areal measurement (i.e., to make an estimate of the activity space for each driver's points for the measured day). These areas, which represent driving "activity space" were used to test for differences (using Student's t-test) in activity space between *SafeBoda* and regular drivers.[78]

Sensitivity analysis: To gauge the robustness of our algorithm in consistently differentiating stops and trips, we performed a sensitivity analysis using 700 runs of the algorithm with different speed and time thresholds/parameters for one randomly selected study participant. Variation in mean number of trip and stop records flagged and mean trip speed across the different parameter combinations was then assessed qualitatively and statistically (using coefficient of variation). Equality of the coefficients of variation in trip speed by parameter combination was assessed using the Feltz and Miller asymptotic test in the *cvequality* R package.[79-81]

Ethical approval and consent: Approval for the study was granted by University of Washington, Makerere University, and the Uganda National Council for Science and Technology. Participants in the study provided verbal consent and were compensated for their time.

RESULTS

From the CAPI questionnaire, the mean age was 32.3 years, mean weekly net income was 25.6 US dollars, and mean years of working in the boda-boda sector was 6.1 years. The majority (88.3%) of the 60 drivers in the study reported previously attending a road safety training. The drivers worked a mean of 12.2 hours and made 17.6 trips per day. *SafeBoda* drivers generally had more years of boda-boda experience, were more educated, and made more money per week (Table 1). Detailed results from the CAPI study are reported elsewhere.[51]

Characteristic	*SafeBoda* drivers N = 30	Regular drivers N = 30
	n (%) or median (IQR)	*n (%) or median (IQR)*
Age (years)	30.5 (8.0)	32.0 (8.8)
Education (years)	9.0 (5.0)	7.0 (4.0)
Weekly net income (in USD)	28.6 (14.3)	17.1 (16.4)
Number of trips per day	20.0 (5.0)	13.5 (10.0)
Hours worked as boda-boda per day	12.0 (2.0)	12.0 (3.0)
Has a driving license	14 (46.7%)	8 (26.7%)
Has other job	13 (43.4%)	10 (33.3%)
Used a phone while driving in past 30 days	7 (23.3%)	18 (60%)
Alcohol use in past 2 hours before driving in past 30 days	0 (0.0%)	5 (41.7%)
Had a road traffic crash in past 6 months	8 (26.7%)	12 (40%)
Own a helmet	30 (100%)	29 (96.7%)
Helmet cost (in USD)	28.6 (0.0)	8.6 (5.5)
Has a reflective jacket	30 (100%)	16 (53.3%)
Has ever received a road safety training	30 (100%)	23 (76.7%)
Has driven on a pedestrian sidewalk in past 30 days	4 (13.3%)	13 (43.3%)
Has driven towards oncoming traffic in past 30 days	2 (6.7%)	12 (40%)
Always used a helmet while driving in past 30 days	30 (100%)	23 (76.7%)
Carried more than one passenger in past 30 days	1 (3.4%)	25 (83.3%)
Years working as a boda-boda driver	5.0 (5.5)	4.0 (3.0)

Note: Continuous variables are reported as median` with interquartile range [IQR] (in parenthesis), while categorical variables are reported as count with percent (in parenthesis). Some observations have missing data.

Table 1 Comparison of baseline characteristics of *SafeBoda* and regular boda-boda drivers in Kampala

Maps of manual versus algorithm processing

Overall, the manual and algorithm approaches appear to have processed the data similarly (i.e., flagged records that were part of a stop or a trip similarly). For instance, of the 88518 GPS records for the participant's data used in the sensitivity analysis, the manual approach identified

26445 points as part of potential trips compared to 20982 points identified by the algorithm. Figures 1 and 2 illustrate how a selection of records from one study participant were processed by the two approaches. Figure 1 shows the unprocessed records with a large cluster of points in the southwest corner of the map. It is not apparent to the eye (i.e., looking at the cluster) which points are associated with each other as a trip or a stop. However, in Figure 2, we can observe the outcome from the manual and algorithm processing. Compared to the algorithm approach, the manual approach seemed to have failed to flag certain clusters of records as part of a stop (circled in red). Moreover, it has removed some data points that were consistent with being part of a trip (circled in green). Conversely, the algorithm approach seemed to have classified some parts of trips as stops (circled in black). With the algorithm, trips of relatively short duration (less than 3 minutes) may have been misclassified as part of a stop.

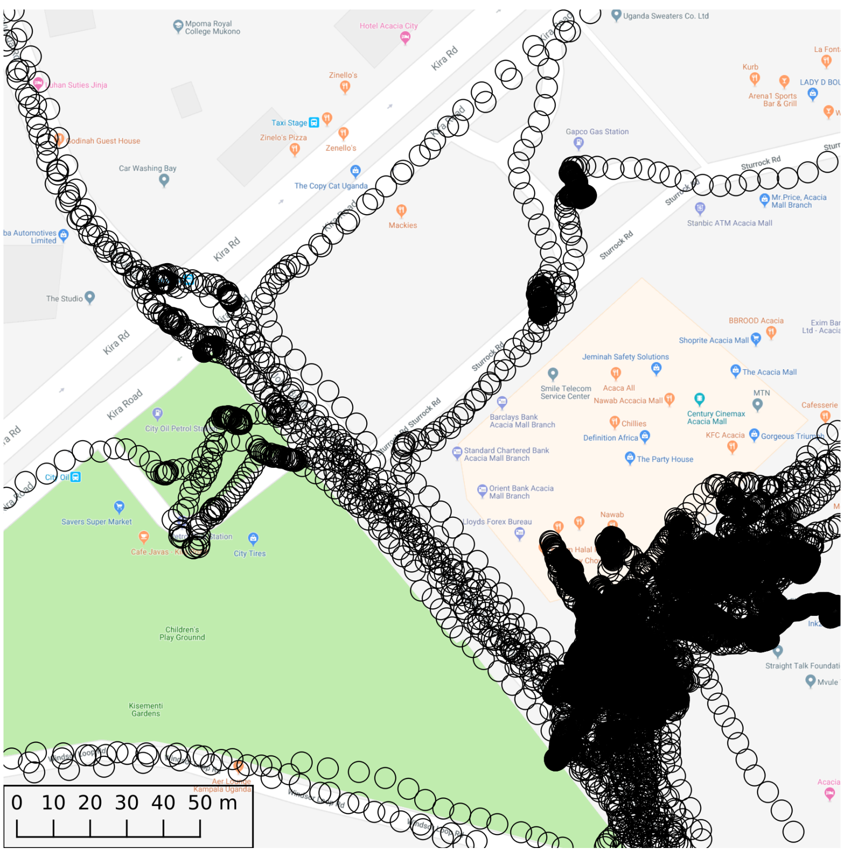


Figure 1: An original (unprocessed) map showing clustered points particularly in the SW area of the map

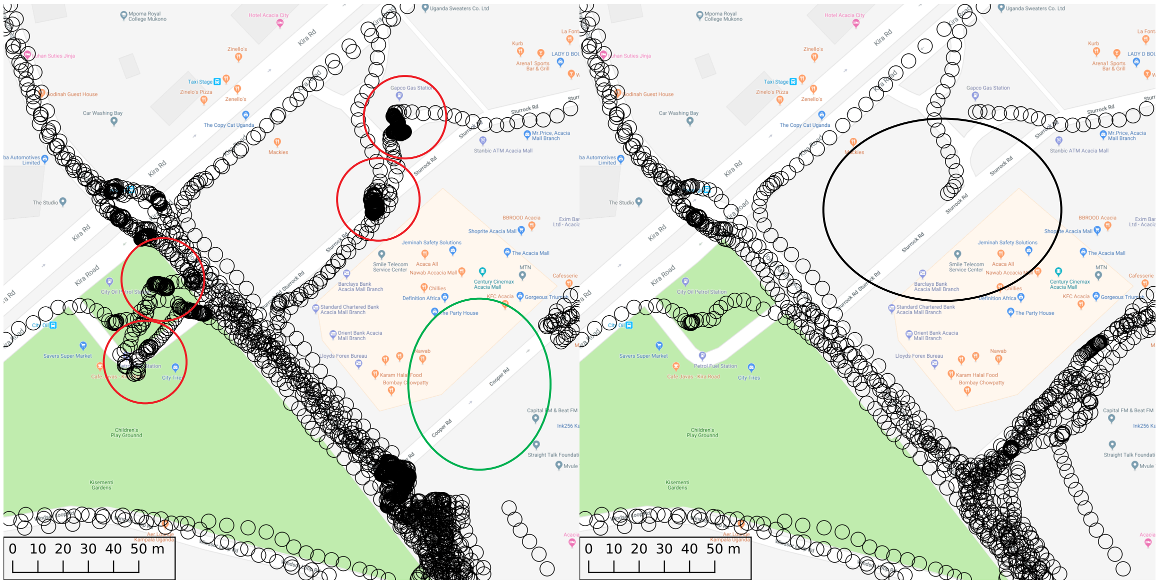


Figure 2: Manually processed (left): note apparent stops that were not manually removed (circled in red), apparent trips that seem to be removed (circled in green); algorithmically processed (right): note apparent trip that was removed (circled in black)

Trip summary findings from the algorithm approach

Results from the GPS data showed that boda-boda drivers in the study made a median of 30.5 trips (IQR: 14.5 trips) per day with a minimum of 1 trip and a maximum of 52 trips. The median trip length was 3.5 km (IQR: 1.1 km) with the shortest trip being 1.1 km and the longest being 6.4 km. A typical trip lasted on average 9.0 minutes (range from 3.0 minutes to 13.8 minutes). The median distance (Euclidean) traveled from the stage was 5.1 km (IQR: 3.3 km) with a minimum and maximum median distance from the stage of 4.2 km and 6.2 km respectively. The median speed on a trip was 22.5 km/h (IQR: 2.2 km/h). A typical boda-boda driver covered a median distance of 115.4 km (IQR: 52.4 km) per day. The median total time spent on trips was 4.9 hours (IQR: 1.8 hours) per day.

On average, *SafeBoda* drivers traveled shorter distances from their stage compared to regular drivers. Conversely, regular drivers made fewer trips per day and covered less distance per day than *SafeBoda* drivers. However overall, driving movements within the city were similar for *SafeBoda* and regular drivers (Table 2).

Characteristic	*SafeBoda* drivers	Regular drivers	P value
Mean trip distance from the stage (km)	4.5 (2.6-6.3)	5.2 (4.4-5.9)	0.20
Minimum trip distance from the stage (km)	3.5 (1.7-5.4)	4.2 (3.4-5.0)	0.24
Maximum trip distance from the stage (km)	5.4 (3.6-7.2)	6.1 (5.4-6.9)	0.19
Mean speed (km/h)	22.8 (21.2-24.4)	22.1 (21.4-22.8)	0.12
Number of trips	32	30	0.46
Mean trip length (km)	3.5 (2.7-4.2)	3.6 (3.3-4.0)	0.41
Mean distance covered per day (km)	110	108	0.85
Number of hours spent driving (hours)	4.7	4.7	0.95
Mean trip duration (minutes)	8.9 (7.3-10.4)	9.4 (8.8-10.1)	0.20

Note: Variables are reported as mean with 95% CIs in parenthesis where applicable

Table 2 Driving behaviors in *SafeBoda* and regular drivers

Heat maps of *SafeBoda* and regular driver routes and trajectories

Qualitatively, the heat maps showed a high concentration of *SafeBoda* trips in the center of the city with most trips around Kampala Central and Makindye Divisions. Trips by regular drivers seem diffused over a wider area (Figure 3). However, our kernel density estimator analysis showed a slighter larger activity space for *SafeBoda* compared to regular drivers although this difference in mean activity space was not statistically significant (Figure 4).

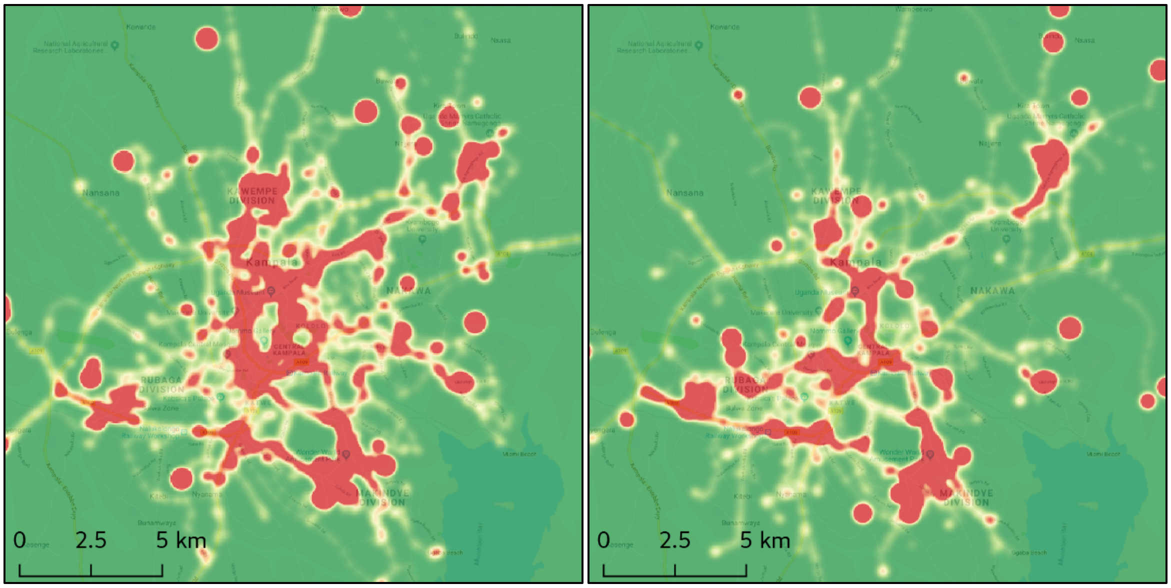


Figure 3: Kernel density estimates heat maps of boda-boda trip trajectories. Left: SafeBoda drivers; right: regular drivers

Activity space analysis

The overall mean of activity space was 5.45 km^2 (SD 5.34), and for *SafeBoda* drivers the mean was slightly larger (6.27 km^2, SD = 5.85), and smaller for regular drivers (4.63 km^2, SD 4.73). A box plot of the activity space area by driver type is shown in Figure 4. Student's t test results indicated no significant difference in activity space between *SafeBoda* and regular drivers (p = 0.24).

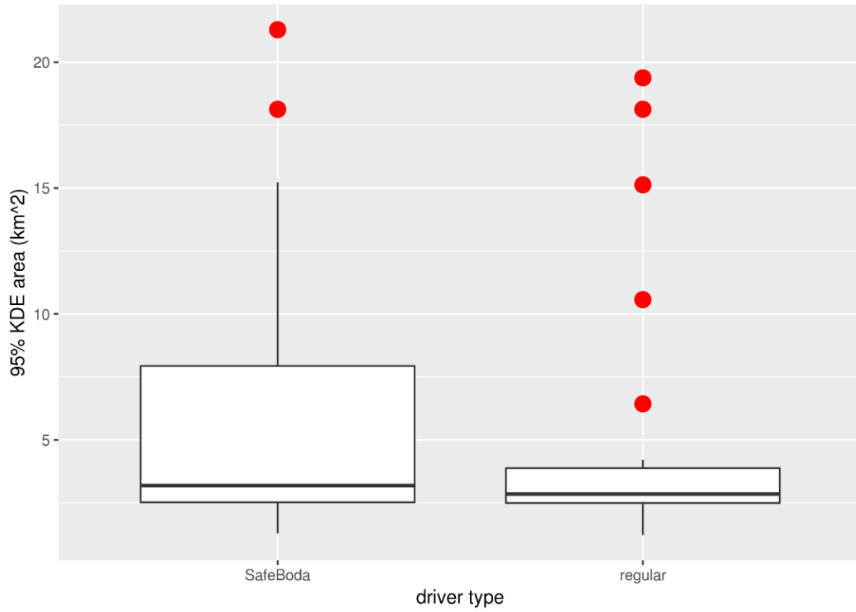


Figure 4: Box plot of activity space for SafeBoda *and regular drivers*

Heat map of speeding hotspots for all boda-boda drivers

The heat map of trip segments where drivers were going at greater than 50 km/h showed scattered speeding hotspots across Kampala (Figure 5). These hotspots seem to generally align with major streets (e.g., Kira road, Kampala road, and Jinja Road) in the city. Although multi-lane, these roads are within built-up areas and drivers are expected to respect the 50 km/h speed limit. There seem to be some locations where *SafeBoda* drivers exceeded the speed limit, but regular drivers did not and vice versa, although it should be acknowledged that their general mobility patterns differ.

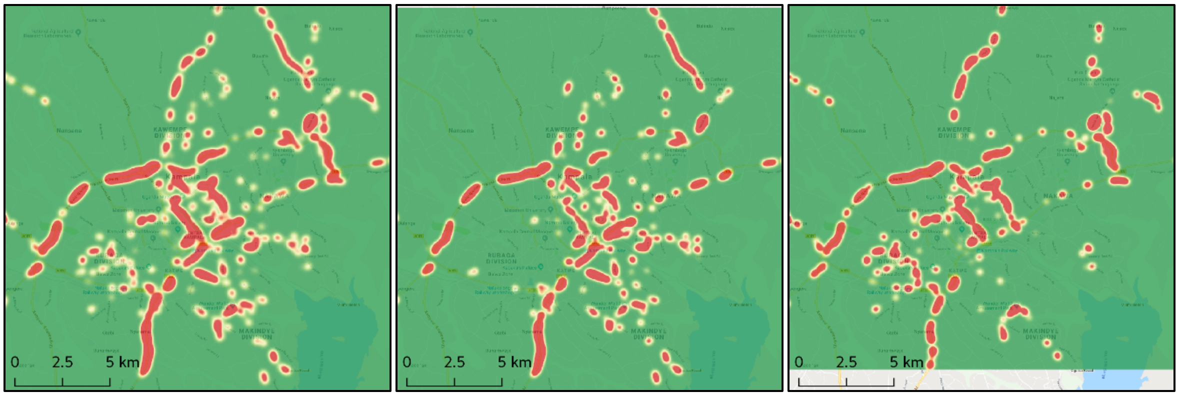


Figure 5: Kernel density estimates heat map of speeding hotspots (i.e., GPS points with speed > 50km/h); left: all data; middle: SafeBoda; right: regular drivers

Sensitivity analysis

Across the 700 parameter combinations used in the sensitivity analysis (only one study participant), the range of mean trip speed was from 21 km/h to 26.3 km/h and the range of time spent in trips was from 2.7 minutes to 6.9 minutes. The mean speed and trip duration (for all study participants) using default parameters (the ones used for the primary analysis) was 22.9 km/h (range from 21.0 km/h to 26.3 km/h) and 9.0 minutes (range from 3.0 minutes to 13.8 minutes).

Qualitatively, mean trip speed did not seem to vary much (median speed of 22.6 km/h; IQR: 1.63 km/h) by parameter combination (Figure 6). In addition, varying one algorithm parameter while holding others constant did not seem to affect mean trip speed (Figures S3-S6). However, the Feltz and Miller asymptotic test showed that the coefficients of variation in trip speed were not equal for each parameter combination ($p<0.001$). The mean coefficient of variation for trip speed across the 700 parameter combinations was 0.52, suggesting a slightly greater level of dispersion around mean trip speed across the different parameter combinations.

While variation in algorithm parameters had little effect on trip speed, we found a marked effect on trip duration (Figure 7 and Figures S7-S10). In terms of number of GPS records identified as part of a trip or a stop, there was some variation by parameter combination with a median of 21712 records (IQR: 2932.3 records) identified as a part of a trip and 66806 records (IQR: 2932.3 records) identified as part of a stop.

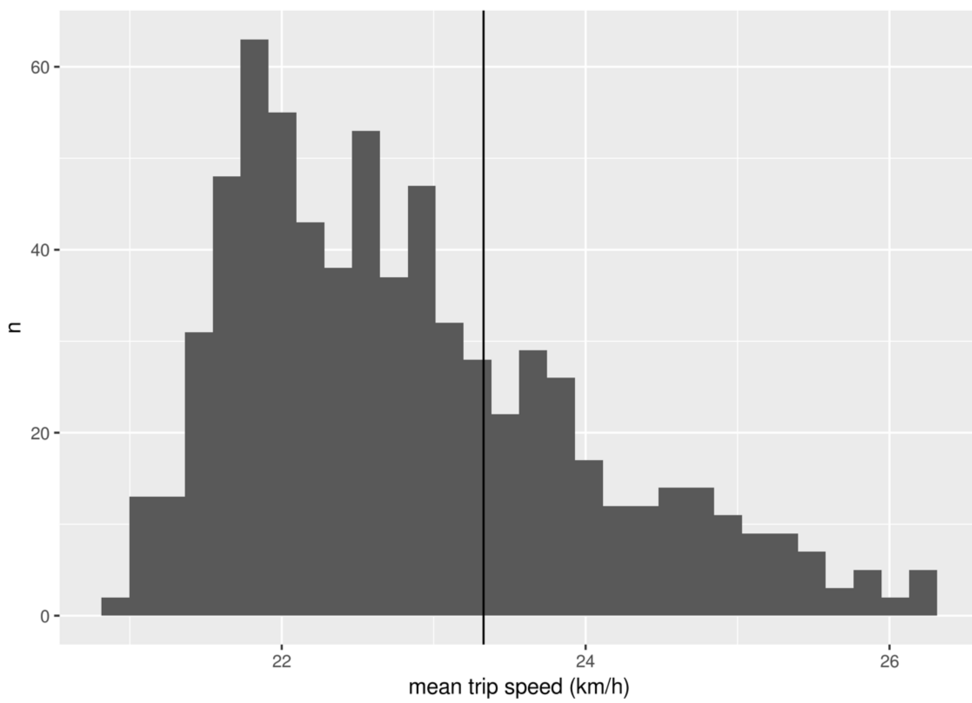


Figure 6: Mean speed across all sensitivity runs (vertical line at selected parameter combination)

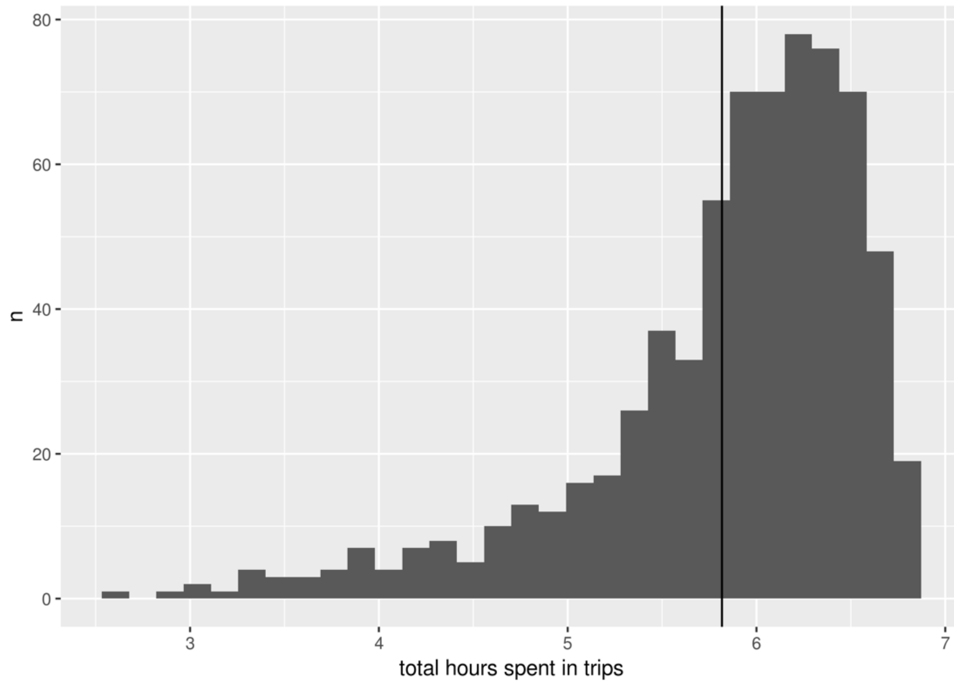


Figure 7: Histogram of total hours spent in trips across sensitivity runs (vertical line at selected parameter combination)

DISCUSSION

In this study, we explored the feasibility of using GPS devices to measure and characterize driving patterns of boda-boda drivers in Kampala. Our findings indicate that GPS tracking in boda-boda drivers is feasible and resulted in high-resolution data that can be used to investigate actual driving behaviors. Of the analysis we conducted, we found that on average boda-boda drivers did not travel more than 7 km from their stage, with trips out of the stage lasting on average 9 minutes. Crude measures such as number of trips taken, distance traveled, and average speed did not differ significantly between *SafeBoda* and regular drivers. Moreover, their heat maps and activity spaces seem to point to shared work environment and passenger destinations. While our sensitivity results showed an estimate for mean trip speed that was robust to changes in the algorithm parameters, the estimated time spent in trips varied across parameter combinations especially when the threshold for what constituted a minimum trip time was large.

Compared to the self-report results from the CAPI questionnaire, boda-boda drivers in the study seemed to have underreported the number of trips they made per day (17.6 in the self-reports versus 31 from the GPS data). Furthermore, while boda-boda drivers self-reported "working" on average 12.2 hours per day, the GPS data showed that only 4.7 hours per day were spent actively driving in trips. The discrepancy in self-reported and algorithm-determined number of trips maybe explained by several factors. For instance, drivers may have understood a trip as starting from their stage, therefore any trips due to trip chaining (i.e., additional trips before return to the stage) would not be reported by the driver but captured by the algorithm. In addition, drivers may have counted as "trips" only those trips that involved carrying a passenger while the algorithm recorded all trips including those without a passenger (e.g., when returning from dropping off a passenger). In essence, the GPS device and self-report may be measuring different things. Secondly, the discrepancy could be due to measurement error especially in the self-reported outcome. One benefit of the GPS device is that it is objective and therefore any systematic errors would be non-differential between the groups being compared.

Our findings collaborate those by Evans et al. in their study of boda-boda drivers in Kampala. We found a similar high concentration of boda-boda trips within central Kampala.[65] However, compared to Evans et al., we found a slightly higher mean trip distance (3.5 km versus 3.04 km) and very different mean speed (22.5 km/h versus 10.0 km/h) and mean trip duration (9.0 minutes vs 18 minutes). These differences could be due to how the two study groups coded trips and analysed their data, however Evans et al. provided limited information in their manuscript on how their measurement and analysis of trip data were done.

This study has some limitations. First, we did not have a gold-standard with which to validate the processing techniques we used in our analysis. For instance, it was not possible to differentiate definitively between stops and trip episodes, especially in situations of short trips or trip chaining with brief stops, such as to pick up or drop off passengers and stops at traffic lights or jams.[82] With the manual approach, there was a potential risk of selection bias from inadvertent removal of actual trip data especially near locations with high concentration of GPS points (e.g., the driver's stage and home). Moreover, at a smaller scale, what appear to be single points may become evident as stops when zoomed to a larger scale. For this reason, the manual editing approach was quite problematic because it required much time spent zooming in and out, and true stops could have easily been missed, leading to misclassification bias. The algorithm approach was also prone to misclassification bias especially in misclassifying shorter trips as stops. In certain situations, the algorithm readily excluded from trips points that appeared to the eye to be part of a moving trip. While there is no reason for us to suspect that any misclassification bias would be different in *SafeBoda* and regular drivers, the overall effect could likely be an underestimation of time spent on trips (and therefore time exposed to a risk of road traffic crash). Although our sensitivity analyses showed that the parameters we selected seemed reasonable for the purposes of this study, selecting an optimal set of parameters for differentiating trips and stops, would be best done using field observations rather than through software simulations. Second, our study was not able to capture hard stops, swerves, very rapid speed changes, and other micro-level risky driving behaviors that would have been better captured with more sensitive devices such as accelerometers. While at the crude level, we did not find evidence of differences in riding patterns between *SafeBoda* and regular drivers, an examination of more precise and targeted measures (e.g., speed through an intersection and riding against traffic) may be needed to determine whether rider safety training results in safer driving among boda-boda drivers. Third, we collected data only for one driving day, which has a few implications: (1) collected data do not necessarily represent typical driving behavior or weather/traffic conditions, (2) subjects may have altered their driving behavior because they knew they were being observed. Lastly, our study did not collect qualitative data to augment the GPS data.

Despite these limitations, our findings suggest that GPS devices can be used to measure boda-boda driver driving patterns in Uganda. Future research in this area should explore using more sensitive devices such as accelerometers to measure finer variations in driving behaviors, consider collecting data from the same participants over several days, and employ a mixed study approach to get both objective GPS data and qualitative data from study participants about their daily experiences as boda-boda drivers in Kampala, as well as their reactions to the use of the GPS devices.[64,65] If combined with qualitative interviews, GPS devices can provide important sources of data on where boda-boda frequently travel in Kampala and how they navigate their work days. This could be useful for road safety planning and enforcement.

Acknowledgments: The authors are grateful to Grace Magambo, Joseph Mugisha, Simon Okurut, and Yefu Chen for their support during the study.

Contributors: All authors conceived the question and study design. KM and OK were involved with data collection. KM and PMH conducted data analysis and wrote the initial draft of the manuscript. All authors were involved with reviewing the manuscript. All authors read and approved the final version of the manuscript for publication.

Funding: This work was supported by the Fogarty International Center of the National Institutes of Health under Grant [#R25 TW009345]. However, the content of this publication is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Competing interests: None. *SafeBoda* was informed about the study but the company had no influence on the study.

Ethics approval: The study was approved by University of Washington IRB, Makerere University School of Public Health IRB, and Uganda National Council for Science and Technology.

Data sharing statement: The data that support the findings of this study are available from the corresponding author, [KM], upon reasonable request.

Chapter 5: Conclusions

In this dissertation, we report findings from five inter-related studies we conducted to assess whether driving behaviors and risk of road traffic crash differed between *SafeBoda* and regular drivers. In the first set of three studies, we found that a lower proportion of *SafeBoda* drivers engaged in driving behaviors associated with increased risk of road traffic crash and injury compared to regular drivers. For instance, compared to regular drivers, helmet use while driving in *SafeBoda* drivers was 21 percent points higher (95% CI: 0.15-0.27; $p<0.001$) based on the CAPI and 45 percent points higher (95% CI: 0.43-0.47; $p<0.001$) based on roadside observation. Furthermore, from the cohort study, we found that *SafeBoda* drivers were 39% less likely to be involved in a RTC than regular drivers after adjusting for age, possession of a driver's license, and education (RR: 0.61, 95% CI: 0.39-0.97, $p=0.04$). Lastly, from the GPS study, we found that GPS devices are acceptable and feasible for measuring boda-boda driver movement in Kampala.

While these were observational studies (with the attendant limitations about making causal inference), we believe our findings have several important implications for the boda-boda sector in Uganda. First, despite almost universal helmet ownership (both self-reported and observed), reported and observed helmet use while driving was found to be considerably lower in regular drivers than in *SafeBoda* drivers. It seems despite owning a helmet, regular drivers seem not interested in wearing it while they are driving. Future studies need to find out why regular drivers are not using their helmets and whether future interventions on helmet use in this group need to focus on awareness campaigns, enforcement, or helmet design. Second, despite the chaos of road traffic in Kampala where even the most careful of drivers is not safe from being involved in a road traffic crash, we were still able to find a large reduced risk of road traffic crash in *SafeBoda* drivers compared to regular drivers. We were not able to determine how the

SafeBoda program was able to exert its effects on safe driving behaviors and risk of road traffic crash. The provision of protective gear and road safety training seemed to have not been the main mediating factors given the universal helmet ownership and null mediation results we found. Two other possible mediators include the strict *SafeBoda* code of conduct and self-selection of road safety-conscious drivers into the *SafeBoda* program. Future studies would need to explore these potential mechanisms in order to identify the components that may be targeted for road safety interventions.

Our studies had several limitations as previously discussed. Some of these limitations can be addressed by additional analyses using the current data or in future studies. For example, while we found a median hospitalization time of 3 days, this is likely an underestimate given that the recorded time was censored at the time of data collection. A future study could consider using survival analysis approaches to provide a more accurate estimate of median hospitalization time. In addition, future studies in motorcycle taxi drivers could include a qualitative component to the GPS study (to understand the driver's experiences with the device) or include new players in the motorcycle taxi sector such as *uberBoda* and *Taxify*.

In conclusion, although there is contradictory evidence in the published literature on the effectiveness of road safety training programs targeting motorcycle drivers, we found evidence that the *SafeBoda* conferred protective effects on its drivers. These effects might not be due to road safety training but due to other components of the *SafeBoda* program such as the code of conduct and its attendant consequences including remedial training, suspension, and dismissal (depending on the nature and context of the offense committed).

References

1. WHO. WHO | Global status report on road safety 2015. *WHO.* 2015.
2. Kardamanidis K, Martiniuk A, Ivers RQ, Stevenson MR, Thistlethwaite K. Motorcycle rider training for the prevention of road traffic crashes. *Cochrane Database Syst Rev.* 2010(10):Cd005240.
3. Adeloye D, Thompson JY, Akanbi MA, et al. The burden of road traffic crashes, injuries and deaths in Africa: a systematic review and meta-analysis. *Bull World Health Organ.* 2016;94(7):510-521a.
4. Kitara D. Boda Boda Injuries in Gulu Regional Hospital, Northern Uganda. *East and Central African Journal of Surgery.* 2011;16(2).
5. Tumwesigye NM, Atuyambe LM, Kobusingye OK. Factors Associated with Injuries among Commercial Motorcyclists: Evidence from a Matched Case Control Study in Kampala City, Uganda. *PLoS One.* 2016;11(2):e0148511.
6. Luchidio M, Kahuthia-Gathu R, Gatebe E. Impact of training boda boda operators and safety status in Kakamega county, Kenya. *International Journal of Advance Research.* 2013;1(9).
7. Raynor B. *Informal transportation in Uganda: a case study of the Boda Boda.* Kenyon College;2014.
8. Oxley J, Yuen J, Ravi MD, et al. Commuter motorcycle crashes in Malaysia: An understanding of contributing factors. *Ann Adv Automot Med.* 2013;57:45-54.
9. Olumide AO, Owoaje ET. Young Age as a Predictor of Poor Road Safety Practices of Commercial Motorcyclists in Oyo State, Nigeria. *Traffic Inj Prev.* 2015;16(7):691-697.
10. UBOS. *Statistical Abstract.* Uganda Bureau of Statistics;2015.
11. Bhalla K, Harrison J, Shahraz S, et al. *Burden of road injuries in sub-Saharan Africa.* Boston, MA: Department of Global Health and Population, Harvard School of Public Health;2013.
12. Ngallaba SE, Majinge C, Gilyoma J, Makerere DJ, Charles E. A retrospective study on the unseen epidemic of road traffic injuries and deaths due to accidents in Mwanza City - Tanzania. *East Afr J Public Health.* 2013;10(2):487-492.
13. PO A, Illika A, Asuzu A. Predictors of road traffic accident, road traffic injury and death among commercial motorcyclists in an urban area of Nigeria. - PubMed - NCBI. 2017.
14. Oginni FO, Ugboko VI, Adewole RA. Knowledge, attitude, and practice of Nigerian commercial motorcyclists in the use of crash helmet and other safety measures. *Traffic Inj Prev.* 2007;8(2):137-141.
15. Johnson O. Prevalence and Pattern of Road Traffic Accidents among Commercial Motorcyclists in a city in Southern Nigeria. (PDF Download Available). *Education Research.* 2012;3(6):537-542.
16. Bachani AM, Hung YW, Mogere S, Akunga D, Nyamari J, Hyder AA. Helmet wearing in Kenya: prevalence, knowledge, attitude, practice and implications. *Public Health.* 2017;144s:S23-s31.
17. Bachani AM, Koradia P, Herbert HK, et al. Road traffic injuries in Kenya: the health burden and risk factors in two districts. *Traffic Inj Prev.* 2012;13 Suppl 1:24-30.
18. Nasong'o WM. *Motorcycle public transport services in Kenya.* University of Nairobi; 2015 2015.
19. Oluwadiya KS, Kolawole IK, Adegbehingbe OO, Olasinde AA, Agodirin O, Uwaezuoke SC. Motorcycle crash characteristics in Nigeria: implication for control. *Accid Anal Prev.* 2009;41(2):294-298.
20. Naddumba E. A cross-sectional retrospective study of Boda boda injuries at Mulago Hospital in Kampala-Uganda. *East and Central African Journal of Surgery.* 2004;9(1).
21. Amoran OE, Eme O, Giwa OA, Gbolahan OB. Road safety practices among commercial motorcyclists in a rural town in Nigeria: implications for health education. *Int Q Community Health Educ.* 2005;24(1):55-64.

22. Nantulya VM, Reich MR. The neglected epidemic: road traffic injuries in developing countries. *Bmj.* 2002;324(7346):1139-1141.
23. Kisaalita WS, Sentongo-Kibalama J. Delivery of urban transport in developing countries: the case for the motorcycle taxi service (boda-boda) operators of Kampala. *Development Southern Africa.* 2007;24(2).
24. de Rome L, Fitzharris M, Baldock M, Fernandes R, Ma A, Brown J. The prevalence of crash risk factors in a population-based study of motorcycle riders. *Injury.* 2016;47(9):2025-2033.
25. McCartt AT, Blanar L, Teoh ER, Strouse LM. Overview of motorcycling in the United States: a national telephone survey. *J Safety Res.* 2011;42(3):177-184.
26. Roehler DR, Naumann RB, Mutatina B, et al. Using baseline and formative evaluation data to inform the Uganda Helmet Vaccine Initiative. *Glob Health Promot.* 2013;20(4 Suppl):37-44.
27. Sebaggala R, Matovu F, Ayebale D, Kisenyi V, Katusiimeh M. *The cost of motorcycle Accidents in Uganda.* 2015.
28. Goodfellow T. Politics on two wheels: motorcycle taxis and divergent development in Rwanda and Uganda. http://africanarguments.org/2015/02/17/politics-on-two-wheels-motorcycle-taxis-and-divergent-development-rwanda-and-uganda-by-tom-goodfellow/. Published 2015. Accessed.
29. Naimanye Andrew, Nganwa Racheal, Kwamusi P, Tomusange M. *Uganda Road Safety Performance Review.* Kampala: Secretariat for the Secretary General's Special Envoy for Road

Safety 2018.
30. SafeBoda. Safety | SafeBoda. https://www.safeboda.com/home. Published 2019. Accessed 01/30/2019, 2019.
31. Mpairwe H. Taxify Go is the new low cost hailing option in Kampala. https://www.techjaja.com/taxify-go-is-the-new-low-cost-hailing-option-in-kampala/. Published 2018. Accessed 8/20, 2018.
32. Simon R. From boda boda rider to business empire; amazing story of Safeboda. https://observer.ug/businessnews/59751-from-boda-boda-rider-to-business-empire-amazing-story-of-safeboda. Published 2019. Accessed 01/30, 2019.
33. Nathan O. Safeboda Reaches 8000 Riders, Rumored to Expand to Three Cities. PC Tech Magazine. https://pctechmag.com/2019/01/safeboda-reaches-8000-riders-rumored-to-expand-to-three-cities/. Published 2019. Updated 2019-01-08. Accessed 01/30, 2019.
34. Woratanarat P, Ingsathit A, Chatchaipan P, Suriyawongpaisal P. Safety riding program and motorcycle-related injuries in Thailand. *Accid Anal Prev.* 2013;58:115-121.
35. Swaddiwudhipong W, Boonmak C, Nguntra P, Mahasakpan P. Effect of motorcycle rider education on changes in risk behaviours and motorcycle-related injuries in rural Thailand. *Trop Med Int Health.* 1998;3(10):767-770.
36. Ivers RQ, Sakashita C, Senserrick T, et al. Does an on-road motorcycle coaching program reduce crashes in novice riders? A randomised control trial. *Accid Anal Prev.* 2016;86:40-46.
37. Galukande M, Jombwe J, Fualal J, Gakwaya A. Boda-boda injuries a health problem and a burden of disease in Uganda: A tertiary hospital survey. *East and Central African Journal of Surgery.* 2009;14(2).
38. Kamulegeya LH, Kizito M, Nassali R, Bagayana S, Elobu AE. The scourge of head injury among commercial motorcycle riders in Kampala; a preventable clinical and public health menace. *Afr Health Sci.* 2015;15(3):1016-1022.
39. UHVI. *Putting helmets on heads – Making roads safe.* 2013.
40. SafeBoda. Safety | SafeBoda. http://www.safeboda.com/home/safety/. Published 2016. Accessed 11/9/2016, 2016.

41. Dean A, Sullivan K, Soe M. Open Source Epidemiologic Statistics for Public Health. In:2013.
42. R Core Team. R: A language and environment for statistical computing. In. Vienna, Austria: R Foundation for Statistical Computing; 2018.
43. Decorps JP. EpiStats: Tools for Epidemiologists. R package version 1.0. In:2017.
44. Ho D, Imai K, King G, Stuart E. MatchIt: Nonparametric preprocessing for parametric causal inference *Journal of Statistical Software.* 2011;42(8):1-28.
45. Edwards JK, Cole SR, Troester MA, Richardson DB. Accounting for Misclassified Outcomes in Binary Regression Models Using Multiple Imputation With Internal Validation Data. *Am J Epidemiol.* 2013;177(9):904-912.
46. Cole SR, Chu H, Greenland S. Multiple-imputation for measurement-error correction. *Int J Epidemiol.* 2006;35(4):1074-1081.
47. van Buuren S, Groothuis-Oudshoor K. mice: Multivariate Imputation by Chained Equations in R *Journal of Statistical Software.* 2011;45(3):1-67.
48. Rubin DB. *Multiple Imputation for Nonresponse in Surveys.* John Wiley & Sons, Inc.; 1987.
49. Wadhwaniya S, Gupta S, Mitra S, et al. A comparison of observed and self-reported helmet use and associated factors among motorcyclists in Hyderabad city, India. *Public Health.* 2017;144s:S62-s69.
50. Bachani AM, Tran NT, Sann S, et al. Helmet use among motorcyclists in Cambodia: a survey of use, knowledge, attitudes, and practices. *Traffic Inj Prev.* 2012;13 Suppl 1:31-36.
51. Muni K, Kobusingye O, Mock C, Hughes JP, Hurvitz PM, Guthrie B. Motorcycle taxi programme increases safe riding behaviours among its drivers in Kampala, Uganda. *Inj Prev.* 2018.
52. Dean A, Sullivan K, Soe M. Open Source Epidemiologic Statistics for Public Health. http://www.openepi.com/. Published 2013. Accessed 8/15, 2015.
53. SMS DUAL. SMS DUAL INFOSOLUTIONS LTD. http://smsdual.info/login.php?from=/index.php? Published 2017. Accessed 12/24, 2017.
54. Carey J V. gee: Generalized Estimation Equation Solver In: Lumley T, Ripley B, eds2015.
55. Dustin T, Teppei Y, Kentaro H, Luke K, Kosuke I. Mediation: R Package for Causal Mediation Analysis. *Journal of Statistical Software.* 2014;59(5).
56. Boele-Vos MJ, de Craen S. A randomized controlled evaluation study of the effects of a one-day advanced rider training course. *Accid Anal Prev.* 2015;79:152-159.
57. Johnson OE, Adebayo A. Effect of Safety Education on Knowledge of and Compliance with Road Safety Signs among Commercial Motorcyclists in Uyo, Southern Nigeria. *Ghana Med J.* 2011;45(3):89-96.
58. Uber. uberBODA has arrived in Kampala. Uber. https://www.uber.com/en-UG/blog/uberboda-kampala/. Published 2018. Accessed 12/3/2018, 2018.
59. Burgess A. *Determining trip and travel mode from GPS and accelerometer data* Indiana: Department of Geography, Indiana University; 2016.
60. Schieltz JM, Okanga S, Allan BF, Rubenstein DI. GPS tracking cattle as a monitoring tool for conservation and management. *African Journal of Range & Forage Science.* 2017;34(3):173-177.
61. Venter C, Joubert J. Using GPS data to characterise driving patterns and fuel use. *Transportation Research Record.* 2013:21.
62. Strauss J, Miranda-Moreno LF, Morency P. Mapping cyclist activity and injury risk in a network combining smartphone GPS data and bicycle counts. *Accid Anal Prev.* 2015;83:132-142.
63. Romanillos G, Moya-Gómez B, Zaltz-Austwick M, Lamíquiz-Daudén PJ. The pulse of the cycling city: visualising Madrid bike share system GPS routes and cycling flow. *Journal of Maps.* 2018;14(1):34-43.
64. Jaligot R, Kemajou A, Chenal J. A new approach to spatial analysis of motorcycle taxis activities – the case of Port-au-Prince, Haiti. *Urban, Planning and Transport Research.* 2017;5(1):78-91.

65. Evans J, O'Brien J, Ng B. Towards a geography of informal transport: Mobility, infrastructure and urban sustainability from the back of a motorbike. *Trans Inst Br Geogr.* 2018;43(4):674-688.
66. Oloo F. Mapping Rural Road Networks from Global Positioning System (GPS) Trajectories of Motorcycle Taxis in Sigomre Area, Siaya County, Kenya. *ISPRS International Journal of Geo-Information.* 2018;7(8):309.
67. Ranacher P, Brunauer R, Trutschnig W, Van der Spek S, Reich S. Why GPS makes distances bigger than they are. *Int J Geogr Inf Sci.* 2016;30(2):316-333.
68. Wu J, Jiang C, Houston D, Baker D, Delfino R. Automated time activity classification based on global positioning system (GPS) tracking data. *Environmental Health.* 2011;10(1):101.
69. Gong H, Chen C, Bialostozky E, Lawson CT. A GPS/GIS method for travel mode detection in New York City. *Computers, Environment and Urban Systems.* 2012;36(2):131-139.
70. Gong L, Morikawa T, Yamamoto T, Sato H. Deriving Personal Trip Data from GPS Data: A Literature Review on the Existing Methodologies. *Procedia - Social and Behavioral Sciences.* 2014;138(Supplement C):557-565.
71. R Core Team. R: A language and environment for statistical computing. In. Vienna, Austria: R Foundation for Statistical Computing; 2017.
72. PostgreSQL. The PostgreSQL Global Development Group. In:2017.
73. PostGIS. The PostGIS Development Group. In:2017.
74. QGIS Development Team. QGIS Geographic Information System. In: Open Source Geospatial Foundation Project; 2019.
75. Wang B, Gao L, Juan Z. A trip detection model for individual smartphone-based GPS records with a novel evaluation method:. *Advances in Mechanical Engineering.* 2017;9(6):1-10.
76. Venables WN, Ripley BD. *Modern Applied Statistics with S.* Fourth ed. New York: Springer; 2002.
77. Evans J. spatialEco. In:2018.
78. Mei-Po K. Interactive geovisualization of activity-travel patterns using three-dimensional geographical information systems: a methodological exploration with a large data set. *Transp Res Pt C-Emerg Technol.* 2000;8(1-6):185–203.
79. Feltz CJ, Miller GE. An asymptotic test for the equality of coefficients of variation from k populations. *Stat Med.* 1996;15(6):646-658.
80. Krishnamoorthy K, Lee M. Improved tests for the equality of normal coefficients of variation. *Computational Statistics.* 2014;29(1):215-232.
81. Marwick B, Krishnamoorthy K. cvequality: Tests for the Equality of Coefficients of Variation from Multiple Groups. In:2019.
82. Schuessler N, Axhausen K. Processing Raw Data from Global Positioning Systems Without Additional Information. *Journal of the Transportation Research Board.* 2009.

Figure S1 Computer-assisted personal interview (CAPI) questionnaire

Study: Motorcycle taxis and road safety in Kampala: comparing health outcomes in SafeBoda and regular motorcycle taxi drivers

Ebibuuzo kunkozesa ya elementi

Record ID ______________________

Interviewer and Study Site Information/Ebikwata ku abuuza ne'kifo

Initials of interviewer/Ammanya g'omunonyereza mu bufunze ______________________

Location of boda-boda stage/Ekifo siteegi ya boda we'sangibwa ______________________
(Describe where the stage is located (include cross streets)/Nyonyola ekifo, ekitundu ne' nguudo eziriranyewo)

Boda-boda stage ID/Endaga kifo ya siteegi ______________________

Name of the part of the city where the stage is located/Erinnya lye'kitundu mu kibuga siteegi weesangibwa ______________________

Number of boda-boda drivers at the stage/Omuwendo gwa'bagoba ba boda ku siteegi ______________________

Number of boda-boda drivers at the stage with visible helmets (either worn or strapped on their motorcycle)/Omuwendo gwa'bagoba/abavuzi ba boda abalina elementi za piki ezirabika (abazambade oba abatazambade ngo'zilabako) ______________________

Other Information (to be completed right before the interview)/Ebintu ebirara ebilina okudibwamu nga omunonyereza tanatandika kunonyereza/kubuliriza

Participant ID/Namba yeyetabye mukunonyereza ______________________

Language questionnaire administered in/Olulimi olubuziibwamu ebibuuzo
◯ English
◯ Luganda

Date of interview/Ennakku zo'mweezi ______________________

Demographic and Other Covariate Information/Ebikwata kumbeera yabantu bomukitundu

Sex of the participant/Ekikula ky'omuntu eyetabyeemu
◯ Female/Mukazi
◯ Male/Musajja

How old are you?/Olina emyaaka emeka? ______________________
(Age in years)

How many years of schooling have you completed? (exclude repeat years)/Wasoma kwenkana ki? ______________________

What is your marital status?/Embeera yo ekwatagana nebyobufumbo eyimiridde etya?	◯ Single, never married/Towasanga ◯ Married or co-habiting/Mufumbo a mubeera mwena ◯ Widowed/Ssemwandu ◯ Separated/Mwayawukana ◯ Divorced/Mwayawukana mu mateeka
Do you have other job (s) apart from working as a boda-boda driver?/Olinayo omulimu/emiriimu emilara okuleka ogwo'buvuzi bwa boda boda?	◯ Yes ◯ No
What other job (s) do you currently have?/Mukiseera kino milimu ki emirara gy'olina?	____________________________
What is your gross average weekly income as a boda-boda driver?/Okutwaliza awamu nga omuvuzi wa boda boda, ofuna sente meka buli sabiiti/wiiki?	____________________________
What is your net average weekly income as a boda-boda driver (that is after subtracting any expenses)?/Okutwaliza awamu, ngomuvuzi wa boda boda ofuna sente meka buli sabiiti/wiiki ngo'jjeko empooza, ne'bilara?	____________________________
What hours do you usually work as a boda-boda?/Otera kukola budde ki obw'olunaku?	◯ Day hours only (6am to 6pm)/Missana ◯ Night hours only (6pm to 6am)/Kiro ◯ Both day and night hours/Missana na Kiro
How many hours per day do you on average work as a boda-boda driver?/Mukugerageranya mulunaku okola essawa mmeka nga omuvuzi wa boda boda?	____________________________ (Enter in hours (no decimals)/Wandiika essawa mubujuvu)
How many trips, on average, do you make per day carrying passengers?/Ovuga luutu z'abasabaze meka mulunaku?	____________________________
Do you drink alcohol? (by alcohol we mean beer, wine, or liquor in a sachet or bottle)/Onnywa omwenge?	◯ Yes ◯ No
Have you driven your motorcycle within 2 hours of drinking alcohol in the past 30 days?/Munaku asaatu eziyise, wali ovuze ku piki mu saawa bbiri oluvanyuma lw'okunywa omwenge?	◯ Yes ◯ No
What type of motorcycle do you have?/E piki piki gyo'lina yakika ki?	◯ Bajaj ◯ Boxer ◯ Senke ◯ Yamaha ◯ Honda ◯ Other
What's the engine size of your motorcycle/Yingini ya piki piki yo yaamanyi ki?	◯ Less than 100cc/Wansi wa cc 100 ◯ 100cc or greater/cc 100 no'kwambuka
Have you used your mobile phone while actively driving your motorcycle in the past 30 days? (excludes use while stopped)/Wali okozeseza ku simu okukuba oba okuwereeza obukaka nga bwovuga munnaku asatu eziyise?	◯ Yes ◯ No
Do you currently own a helmet for your motorcycle?/Olina elementi ey'obwannanyini?	◯ Yes ◯ No

What's the type of your helmet?/Elementi yo ya kikaa ki?

____
(Write the brand name and then the type in parenthesis/Wandiika omukozi wa elementi)

What's the color of your helmet?/Elementi yo yaalangi ki?

____

How long have you had your helmet?/Elementi yo omaze nayo banga ki?

____

How much (in UGX) did you pay for your helmet? (applicable only to those who purchased their helmet, not given for free)/Elementi yo wagigula ssente mmeka eza Uganda?

____

Do you currently own any of these protective gear to wear while driving your motorcycle? (mark all that apply)/Olina yo kubyokwekumisa bino wammanga byoosobola okwambala ngo'vuga piki piki?

☐ Sturdy boots/Engatto ezivuga
☐ Gloves/Gilaavuzi
☐ Reflective jacket/Ekkoti eyaka/jaketi
☐ Helmet/Elementi
☐ Protection glasses/Ggalubindi
☐ Thick pants/trousers/Empale engumu
☐ None/Tewali

Do you own the motorcycle you use for boda-boda?/Piki piki gyovuga yiyo?

◯ Yes
◯ No

Are you a member of any of these boda-boda programs? (mark all that apply)/Oliko mu nteka teka/pulogulaamu yonna eyabavuzi ba boda boda?

☐ SafeBoda
☐ BodaPro
☐ Tugende
☐ None/Tewali

Have you ever received a road safety training from programs such as Uganda Helmet Vaccine Initiative, SafeBoda, BodaPro, or Tugende?/Wali otendekeddwako kunkozesa yolugudo okuva mu bibiina nga Uganda Helment Vaccine initiative, Safeboda, Bodapro oba Tugende?

◯ Yes
◯ No

How long (in years) have you worked as a boda-boda driver?/Omulimu gwo'kuvuga boda ogumazeemu banga ki?

____

Are you a member of any boda-boda union or association?/Olina ekibiina kyona ekyabavuzi ba boda boda kyolimu?

◯ Yes
◯ No

What's the name of your union or association?/Ekibiina bakiyita batya?

____

Do you have a driver's license?/Olina ebisanyizo ebikukiriza okuvuga boda?

◯ Yes
◯ No

How did you learn how to drive a motorcycle?/Wayiga otya okuvuga pikipiki?

◯ Self taught (received no help from other people)/Weka
◯ Friend/relative/peer/Mukwano/ owo'luganda
◯ Driving school/Muttendekero lya'bavuzi
◯ Other/Ewalala wonna

At what age did you start driving a motorcycle?/Watandika okuvuga piki ngo'lina emyaka emeka?

____

How long (in days) did it take you from when you started driving a motorcycle to when you started driving it on public roads?/Kyakutwalira banga ki okuva bwewatandiika okuyiga okutuusa lwewatandika okuvugira ku nguudo ezikozesebwa abantu?

____

How long (in weeks) did it take you from when you started driving a motorcycle to when you started carrying passengers as a boda-boda driver?/Kyakutwalira ssabiiti mmeka okuva lwewatandiika okuyiga paka bwe watandiika okutwala abantu nga omuvuzi wa boda boda?

____

Were you involved in a road traffic crash in the past 6 months (by crash we mean an incident where you collided with another vehicle (e.g. car, motorcycle), a pedestrian, an animal, or an object)/Wali ofunye ku kabenje akekika kyonna mumyeezi 6 egiyise? (akabenje tutegeeza wewekoona nekidukka ekirala, nga emotoka,pikipki, omusabaze,ensolo oba ekintu kyonna).

◯ Yes
◯ No

Who were the other parties involved in your most recent crash? (mark all that apply)/B'ani abalala abali mu kabenje ako?

☐ Motorcycle/Piki piki
☐ Car/truck/Mmotoka
☐ Bus/Baasi
☐ Bicycle/Akagaali
☐ Pedestrian/Omutanbuuze
☐ Animal/Ensolo
☐ Tree/pole/road barriers/Omuuti/Ekikondo
☐ None/Tewali

Were you wearing a helmet at the time of your most recent crash?/Wali oyanbade elementi mu kabenje kewakasembayo okufuna?

◯ Yes
◯ No

At the time of your most recent crash, which of the following was true of you? (mark all that apply)/Mukiseera kyewafuniramu akabenje, kiki ku bino Kyandiiba nga kye'kyakuleetera akabenje?

☐ Speeding/Ndiima
☐ Faulty brakes/Ebiziyiza ebifu
☐ Making an illegal turn/Okukyukira awakyamu
☐ Slippery surface/Obuseerezi
☐ Drink driving/Obutamiivu
☐ Running a red light or stop sign/Obutagondera bitaala
☐ Wrong way driving/Okuvugira ku ludda olukyaamu
☐ Distracted driving/Obutassayo Mwoyo
☐ Other/Ekirara kyonna

Who do you think was at fault for your most recent crash?/Ani gwo'lowooza eyali mu nsobi mu kabenje kewasembayo okufuna?

◯ Myself/Nze mwennyini
◯ The other party (ies) involved/Abalala abali mukabenje
◯ No one/Tewali
◯ Not sure/Simanyi bulungi

Where did your most recent crash occur?/Akabenje kewasembayo okufuna wakafunira luddawa?

____
(Road or street name where the crash happened (ask for cross streets or a landmark if possible)/Erinnya lyo'lugudo awaagwa akabenje))

When did your most recent crash occur?/Wafuna ddi akabenje ako?

◯ Morning (6am-12pm)/Kumakya
◯ Afternoon (12pm-6pm)/Mu ttuntu
◯ Night (6pm-6am)/Ekiro

Did you sustain an injury from your most recent crash?/Wafuna mu ebisago okuva ku kabenje kewakasembayo okufuna?

◯ Yes
◯ No

What body part (s) did you injure? apply)/Bitundu ki kumubiri ebyakosebwa?

☐ Head/Mutwe (mark all that
☐ Neck/Nsiingo
☐ Arms/hands/Mikono/engalo
☐ Chest/Kifuba
☐ Legs/Magulu
☐ Pelvis/Enkoggo
☐ Feet/Bigere
☐ Spinal cord/Enkiizi
☐ Other/Ekirara kyonna

Did your injury require a visit to a health facility to get care?/Oluvanyuma lwa'kabenje, kyali kikwetagisa okufuna obujanjabi okuva mu ddwaliro?

◯ Yes
◯ No

Did your injury require an admission (that is inpatient care)?/Obuvune bwo bwali bwetagisa okuweebwa ekitanda?

◯ Yes
◯ No

For how many days were you admitted at the healthcare facility for your injury?/Wamala ennaku meka ku kitanda?

____

Have you driven your motorcycle on a curb or pavement in the past 30 days?/Wali ovugiddeko piki piki yo ku peevumenti munnaku 30 eziyise?

◯ Yes
◯ No

Have you driven your motorcycle in the wrong direction (opposite to oncoming traffic) in the past 30 days?/Wali ovugiddeko piki piki yo ku ludda olukyamu munnaku 30 eziyise?

◯ Yes
◯ No

Have you carried more than one passenger per trip in the past 30 days?/Wali oweeseko ku basabaze abassukka mu omu munnaku 30 eziyise?

◯ Yes
◯ No

Outcome Information

(all questions refer to while you are working as a boda-boda driver/Bino ebibuzo byonna bikwatagana ku banga/obudde nga okola nga omuvuzi wa boda boda)

Do you ever wear a helmet while driving a motorcycle?/Otera okwambala elementi nga ovuga piki piki?

◯ Yes
◯ No

In the past 30 days, how often would you say you wore a helmet while driving a motorcycle?/Munnaku asaatu eziyise, oyambaliddeme elementi emirundi emmeka?

◯ Never/Tewali mulundi nagumu
◯ Sometimes/Emirundi egiimu
◯ Always (all the time)/Emirundi gyonna
◯ Don't know (unsure)/Simanyi
◯ Refused to answer/Sijja kuddamu
◯ Do not own a helmet/Sirina elementi

Exposure Information (to be completed by the interviewer)/byakujjuzibwaamu oyo abuuza

Is the subject wearing anything to identify him as a SafeBoda driver?/Eyetabye mukunonyereza alika ekintu kyonna kya'yamadde ekiraga nti muvuzi wa Safe boda?	◯ Yes ◯ No
If Yes, which one? (mark all that apply)/Bwekibanga kituufu, kinyonyole?	☐ Sticker/Sitiika ☐ Jacket/Jaketi ☐ Other/Ekirara kyonna

Exposure Information Questions for the Participant/Byakujjuzibwa abuuziibwa

Were you ever a SafeBoda driver?/Wali abaddeko mu kibiina kta Safe Boda?	◯ Yes ◯ No
Are you currently a driver for the SafeBoda program?/Oli muvuzi wa Safeboda kati?	◯ Yes ◯ No
When did you become a SafeBoda driver?/Wafuuka ddi omuvuzi wa Safeboda?	______________________
Do you currently use the SafeBoda app for accepting passenger request for rides?/Mukiseera kino okozesa kunkola ya Safeboda ey'okumutimbagano okufuna abasabaze?	◯ Yes ◯ No
How often would you say you wore a helmet while a motorcycle before you became a SafeBoda driver?/Mirundi gyenkanaki gyo'yinza okugamba nti wayambala nga elementi nga tonafuuka muvuzi wa Safeboda?	◯ Never/Tewali mulundi nagumu driving ◯ Sometimes/Emirundi egiimu ◯ Always (all the time)/Emirund byonna ◯ Don't know (unsure)/Simanyi ◯ Refused to answer/Sijja kuddamu ◯ Did not own a helmet/Sirina elementi
Do you have a friend or a peer who is a SafeBoda driver?/Olinayo mukwano gwo yenna omuvuzi wa Safeboda?	◯ Yes ◯ No
Has having a friend or peer who is a SafeBoda driver impacted your use of any of the following? (mark all that apply)/Okuba ne mukwano gwo omuvuzi wa Safeboda, kikoseza kitya enkozessayo eyabino wammanga?	☐ Helmet/Elementi ☐ Reflective jacket/Jaketi ☐ Not at all/Tewali yadde

Cohort Study of RTC and RTI Incidence Consent Information

(Please read the cohort study consent form before completing this section)

Participant consented to participate in cohort study/Eyetabyeemu akkiriza okwetaba mu musomo	◯ Yes ◯ No
Primary mobile phone number/Ennamba ye'ssimu esooka	______________________
Secondary mobile phone number/Ennamba ye'ssimu eyokubiri	______________________
Date of cohort enrollment/Ennaku zo'mwezi omuntu mwa'londebbwa	______________________ (Enter date of cohort enrollment here/Yingiiza ennaku zo'mwezi wano)

Riding Behavior Study Consent Information

(Please complete this section after receipt of verbal consent from the participant for the cohort study/Juuza bino oluvanyuma lwo'mubuziibwa okukkiriza okwetaba mu kusoma kuno).

Participant consented to participate in pilot study to measure driving patterns among boda-boda drivers Kampala/Omubuziibwa akkiriza okwetaba mu kusoma kunvuga ya boda mu bavuzi be'Kampala.	◯ Yes ◯ No in
Date of GPS tracker and accelerometer installation/Ennaku zo'mwezi obuuma bu kalimageezi lwebunaatekebwako	____________________________ (Can be scheduled in advance/Esobola okuteesebwako nga bukyaali)
GPS tracker ID/Enyinyonyola ya'kuuma akalaga ekifo	____________________________
Accelerometer ID/Enyinyonyola ya'kuuma akapiima omuliiro	____________________________

Figure S2 Directed acyclic graph of the relationship between *SafeBoda* and helmet use

This figure shows variable relationships (based on theory and existing injury literature) among the exposure (*SafeBoda* program), primary outcome (self-reported helmet use), and several covariates (e.g., age, education, and past crash history).

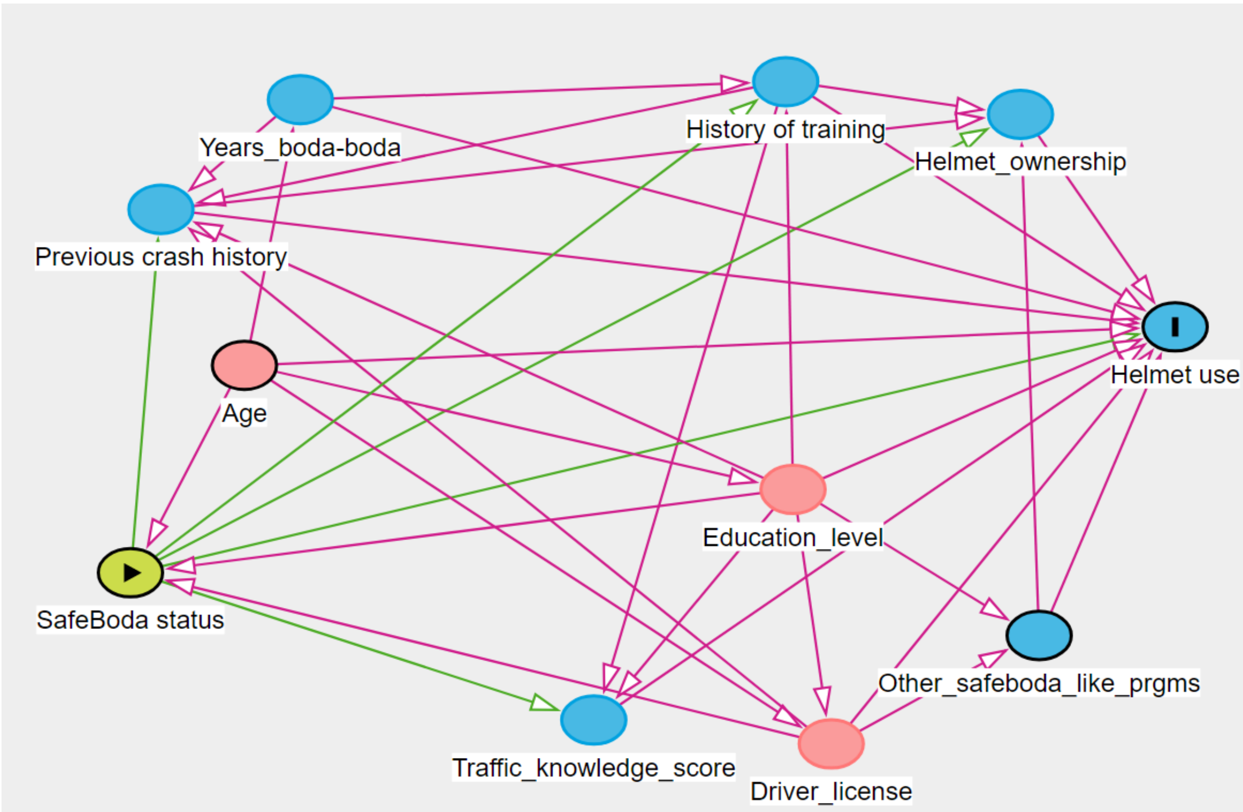


Figure S3 Summary of matched sets

```
Call:
matchit(formula = safeboda ~ education + age + boda_rtc + license,
    data = df2, method = "nearest", distance = "logit", caliper = 0.05)

Summary of balance for all data:
           Means Treated Means Control SD Control Mean Diff eQQ Med eQQ Mean eQQ Max
distance          0.5566        0.4301     0.1645    0.1265  0.0808   0.1277  0.2849
education         8.4175        7.9550     3.6011    0.4625  0.0000   0.5567  2.0000
age              33.4897       32.1100     6.9837    1.3797  1.0000   1.4897  7.0000
boda_rtc0         0.7938        0.6650     0.4732    0.1288  0.0000   0.1289  1.0000
boda_rtc1         0.2062        0.3350     0.4732   -0.1288  0.0000   0.1289  1.0000
licenseYes        0.6701        0.3350     0.4732    0.3351  0.0000   0.3351  1.0000


Summary of balance for matched data:
           Means Treated Means Control SD Control Mean Diff eQQ Med eQQ Mean eQQ Max
distance          0.4940        0.4906     0.1710    0.0034  0.0046   0.0044  0.0089
education         8.4496        8.2636     3.4245    0.1860  1.0000   0.5736  2.0000
age              32.9225       32.0698     6.6523    0.8527  1.0000   1.1318  6.0000
boda_rtc0         0.7132        0.7209     0.4503   -0.0078  0.0000   0.0078  1.0000
boda_rtc1         0.2868        0.2791     0.4503    0.0078  0.0000   0.0078  1.0000
licenseYes        0.5039        0.5039     0.5019    0.0000  0.0000   0.0000  0.0000

Percent Balance Improvement:
           Mean Diff. eQQ Med eQQ Mean  eQQ Max
distance      97.3083 94.3363  96.5238  96.8935
education     59.7760    -Inf  -3.0434   0.0000
age           38.1953  0.0000  24.0256  14.2857
boda_rtc0     93.9821  0.0000  93.9845   0.0000
boda_rtc1     93.9821  0.0000  93.9845   0.0000
licenseYes   100.0000  0.0000 100.0000 100.0000

Sample sizes:
          Control Treated
All           200     194
Matched       129     129
Unmatched      71      65
Discarded       0       0
```

Figure S4 Histograms showing distribution of propensity scores

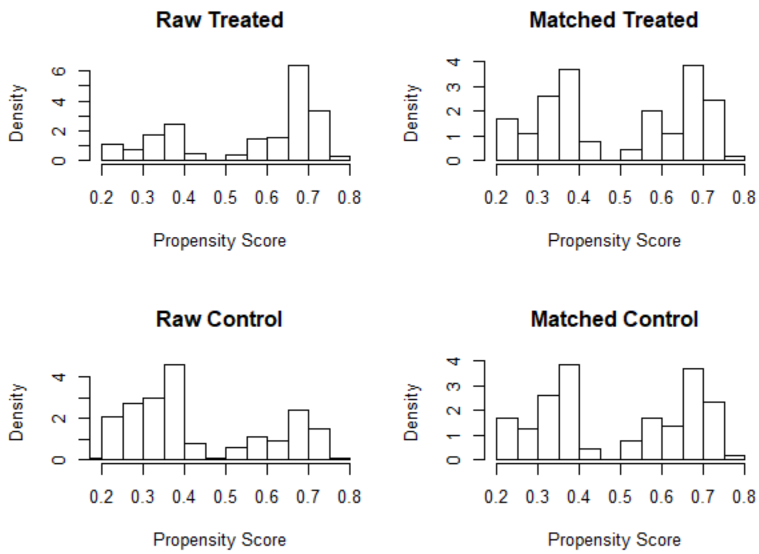
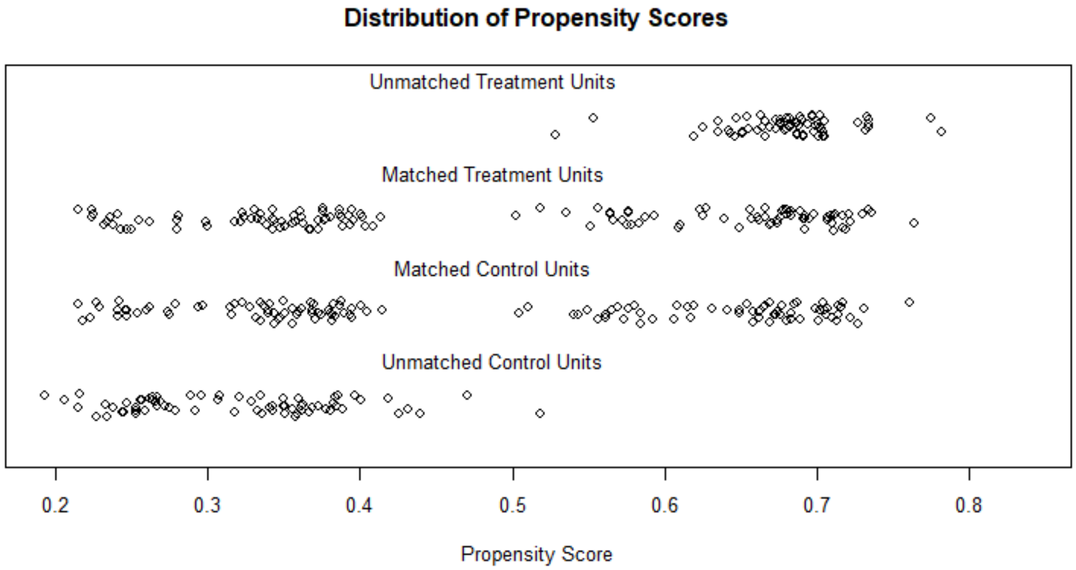


Figure S5 Summary of distribution of propensity scores using “jitter”



Supplemental online materials

Figure S1 Computer-assisted personal interview questionnaire

Cohort baseline questionnaire

Record ID ______________________

Interviewer and Study Site Information/Ebikwata ku munonyereza nekifo mwebanonyereza (Use this questionnaire for only cohort participants not recruited from the Cross-sectional Study of Boda-Boda Drivers/Kozesa ebibuuzo bino ku bantu abo bokka bemutawandisa kwetaba mu kunonyereza kuno)

Initials of interviewer/Ammanya g'omunonyereza mu bufunze ______________________

Location of boda-boda stage/Ekifo sitegi ya boda we'sangibwa ______________________
(Describe where the stage is located (include cross streets)/Nyonyola ekifo, ekitundu ne'nguudo eziriraanyewo)

Boda-boda stage ID/Endaga kifo ya siteegi ______________________

Name of the part of the city where the stage is located/Erinya lye'kitundu mu kibuga siteegi weesangiibwa ______________________

Number of boda-boda drivers at the stage/Omuwendo gw'abagoba ba boda ku sitegi ______________________

Number of boda-boda drivers at the stage with visible helmets (either worn or strapped on their motorcycle)/Omuwendo gwabagoba/abavuzi ba boda abalina elementi za piki piki ezirabika (abazambadde oba abatazambade nga ozilabako) ______________________

Other Information (to be completed right before the interview)/Ebintu ebirara ebilina okudibwamu nga omunonyereza tanatandika kunonyereza/kubuliriza

Participant ID/Ennamba yooyo eyetabye mukunonyereza ______________________

Language questionnaire administered in/Olulimi olubuzibwamu
- ◯ English
- ◯ Luganda

Date of interview/Ennakku zo'mweezi ______________________

Demographic and Other Covariate Information/Ebikwata kumbeera y'abantu b'omukitundu

Sex of the participant/Ekikula ky'omuntu eyetabyeemu
- ◯ Female/Mukazi
- ◯ Male/Musajja

How old are you?/Olina emyaka emeka ______________________
(Age in years/Emyaka)

How many years of schooling have you completed? (exclude repeat years)/Wasoma kwenkana ki? ______________________

What is your marital status?/Embeera yo ekwatagana nebyobufumbo eyimiridde etya?	◯ Single, never married/Towasanga ◯ Married or co-habiting/Mufumbo a mubeera mwena ◯ Widowed/ Ssemwandu ◯ Separated/Mwayawukana ◯ Divorced/Mwayawukana mu mateeka
Do you have other job (s) apart from working as a boda-boda driver?/Olinayo omulimu/emiriimu emirara okuleka ogwo'buvuzi bwa boda boda?	◯ Yes ◯ No
What other job (s) do you currently have?/Mukiseera kino milimu ki emirara gy'olina?	________________________________
What is your gross average weekly income as a boda-boda driver?/Okutwaliza awamu nga omuvuzi wa boda boda, ofuna sente meka buli sabiiti/wiiki?	________________________________
What is your net average weekly income as a boda-boda driver (that is after subtracting all expenses)?/Okutwaliza awamu, ngomuvuzi wa boda boda ofuna sente meka buli sabiiti/wiiki ngo'jjeko empooza, ne'bilara?	________________________________
What hours do you usually work as a boda-boda?/Otera kola budde ki nga omuvuzi wa boda boda?	◯ Day hours only (6am to 6pm)/Missana ◯ Night hours only (6pm-6am)/Kiro ◯ Both day and night hours/Missana na Kiro
How many hours per day do you on average work as a boda-boda driver?/Mukugerageranya mulunaku olumu, okola essawa mmeka nga omuvuzi wa boda boda?	________________________________ (Enter in hours (no decimals)/wandiika essawa mubujuvu)
How many trips, on average, do you make per day carrying passengers?/Mukugerageranya ovuga luutu z'abasabaze meka mulunaku?	________________________________
Do you drink alcohol? (by alcohol we mean beer, wine, or liquor in a sachet or bottle)/Onnywa omwenge?	◯ Yes ◯ No
Have you driven your motorcycle within 2 hours of drinking alcohol in the past 30 days?/Munaku asaatu eziyise, wali ovuze ku piki mu saawa bbiri oluvanyuma lw'okunywa omwenge?	◯ Yes ◯ No
What type of motorcycle do you have?/Kikaaki ekya piki piki ky'olina?	◯ Bajaj ◯ Boxer ◯ Senke ◯ Yamaha ◯ Honda ◯ Other
What's the engine size of your motorcycle?/Yingini ya piki piki yo yaamanyi ki?	◯ Less than 100cc/Wansi wa cc 100 ◯ 100cc or greater/cc 100 no'kwambuka
Have you used your mobile phone while actively driving your motorcycle in the past 30 days? (excludes use while stopped)/Wali okozeseza ku simu yo nga ovuga boda boda yo mu bbanga ely'enaku assatu eziyise?	◯ Yes ◯ No

Do you currently own a helmet for your motorcycle?/Olina elementi ey'obwannanyini ku piki piki yo?	◯ Yes ◯ No
What's the type of your helmet?/Elementi yo ya kikaaki?	______________________________ (Write the brand name and then the type in parenthesis/wandiika errinya nekika kya elementi)
What's the color of your helmet?/Elementi yo yaalangi ki?	______________________________
How long have you had your helmet?/ Elementi yo omaze nayo banga ki?	______________________________
How much (in UGX) did you pay for your helmet? (applicable only to those who purchased their helmet, not given for free)/Elementi yo wagigula ssente mmeka eza Uganda?	______________________________
Do you currently own any of these protective gear to wear while driving your motorcycle? (mark all that apply)/Mukiseera kino olinayo kub'yokwekumisa bino w'amanga eby'okwambala ngo'vuga piki piki?	☐ Sturdy boots/Engatto ezivuga ☐ Gloves/Gilaavuzi ☐ Reflective jacket/Ekkoti eyaka/jaketi ☐ Helmet/Elementi ☐ Protection glasses/Ggalubindi ☐ Thick pants/trousers/Empale engumu ☐ None/Tewali
Do you own the motorcycle you use for boda-boda?/Piki piki gy'ovuga yiyo?	◯ Yes ◯ No
Are you a member of any of these boda-boda programs? (mark all that apply)/Olina pulogulaamu yonna eyabavuzi ba boda boda gy'olimu nga omu ku bbo/member?	☐ SafeBoda ☐ BodaPro ☐ Tugende ☐ None/Tewali
Have you ever received a road safety training from programs such as Uganda Helmet Vaccine Initiative, SafeBoda, BodaPro, or Tugende?/Wali otendekeddwako kunkozesa y'olugudo ennungi okuva mu bibiina nga Uganda Helment Vaccine initiative, Safe boda, Bodapro oba Tugende?	◯ Yes ◯ No
How long (in years) have you worked as a boda-boda driver?/Omulimu gwo'kuvuga boda ogumazeemu banga ki?	______________________________
Are you a member of any boda-boda union or association?/Olina ekibiina kyonna ekyabavuzi ba boda boda mwooli?	◯ Yes ◯ No
What's the name of your union or association?/Ekibiina kyamwe bakiyita batya?	______________________________
Do you have a driver's license?/Olina ebisanyizo ebikukiriza okuvuga boda?	◯ Yes ◯ No
How did you learn how to drive a motorcycle?/Wayiga otya okuvuga piki piki?	◯ Self taught (received no help from other people)/Weka ◯ Friend/relative/peer/Mukwano/ owo'luganda ◯ Driving school/Muttendekero lya'bavuzi ◯ Other/Ewalala wonna
At what age did you start driving a motorcycle?/Watandika okuvuga piki ngo'lina emyaka emeka?	______________________________

How long (in days) did it take you from when you started driving a motorcycle to when you started driving it on public roads?/Kyakutwalira ennaku meka okuva bwewatandiika okuyiga okutuusa lwewatandika okuvugira ku nguudo ezikozesebwa abantu bonna?

____

How long (in weeks) did it take you from when you started driving a motorcycle to when you started carrying passengers as a boda-boda driver?/Kyakutwalira ssabiiti mmeka okuva lwewatandiika okuvuga pikipiki paka lwewatandika okutwaala abasabaze nga omuvuzi wa boda boda?

____

Were you involved in a road traffic crash in the past 6 months (by crash we mean an incident where you collided with another vehicle (e.g. car, motorcycle), a pedestrian, an animal, or an object)/Wali ofunye ku kabenje kona mumyeezi 6 egiyise? (akabenje tutegeeza wewekoona nekidukka ekirala, nga emotoka,pikipki, omusabaze,ensolo oba ekintu kyonna).

◯ Yes
◯ No

Who were the other parties involved in your most recent crash? (mark all that apply)/B'ani abalala abali mu kabenje ako?

☐ Motorcycle/Piki piki
☐ Car/truck/Mmotoka
☐ Bus/Baasi
☐ Bicycle/Akagaali
☐ Pedestrian/Omutanbuuze
☐ Animal/Ensolo
☐ Tree/pole/road barriers/Omuuti/Ekikondo
☐ None/Tewali

Were you wearing a helmet at the time of your most recent crash?/Wali oyambadde elementi mu kabenje kewakasembayo okufuna?

◯ Yes
◯ No

At the time of your most recent crash, which of the following was true of you? (mark all that apply)/Mukiseera kyewafuniramu akabenje, kiki ku bino Kyandiiba nga Kye'kyakuleetera akabenje?

☐ Speeding/Ndiima
☐ Faulty brakes/Ebiziyiza ebifu
☐ Making an illegal turn/Okukyukira awakyamu
☐ Slippery surface/Obuseerezi
☐ Drink driving/ Obutamiivu
☐ Running a red light or stop sign/Obutagondera bitaala
☐ Wrong way driving/Okuvugira ku ludda olukyaamu
☐ Distracted driving/Obutassayo Mwoyo
☐ Other/Ekirara kyonna

Who do you think was at fault for your most recent crash?/Ani gwo'lowooza eyali mu nsobi mu kabenje kewasembayo okufuna?

◯ Myself/Nze mwennyini
◯ The other party (ies) involved/Abalala abali mukabenje
◯ No one/Tewali
◯ Not sure/Simanyi bulungi

Where did your most recent crash occur?/Akabenje kewasembayo okufuna wakafunira luddawa?

____
(Road or street name where the crash happened (ask for cross streets or a landmark if possible)/erinnya lyo'lugudo awaagwa akabenje)

When did your most recent crash occur?/Akabenje kewasembayo okufuna kaliwo ddi?

◯ Morning (6am-12pm)/Kumakya
◯ Afternoon (12pm-6pm)/Mu ttuntu
◯ Night (6pm-6am)/Ekiro

Did you sustain an injury from your most recent crash?/Wafuna mu ebisago okuva ku kabenje kewakasembayo okufuna?	◯ Yes ◯ No
What body part (s) did you injure? (mark all that apply)/Bitundu ki eby'omubiri ebyakosebwa?	☐ Head/Mutwe ☐ Neck/Nsiingo ☐ Arms/hands/Mikono/ engalo ☐ Chest/Kifuba ☐ Legs/Magulu ☐ Pelvis/Enkoggo ☐ Feet/Bigere ☐ Spinal cord/ Enkiizi ☐ Other/Ekirara kyonna
Did your injury require a visit to a health facility to get care?/Oluvanyuma lwa'kabenje, kyali kikwetagisa okufuna obujanjabi okuva mu ddwaliro?	◯ Yes ◯ No
Did your injury require an admission (that's inpatient care)?/Obuvune bwo bwali bwetagisa okuweebwa ekitanda?	◯ Yes ◯ No
For how many days were you admitted at the healthcare facility for your injury?/Mamala ennaku meka ku kitanda?	________________________________
Have you driven your motorcycle on a curb or pavement in the past 30 days?/Wali ovugiddeko piki piki yo ku peevumenti munnaku 30 eziyise?	◯ Yes ◯ No
Have you driven your motorcycle in the wrong direction (opposite to oncoming traffic) in the past 30 days?/Wali ovugiddeko piki piki yo ku ludda olukyamu munnaku 30 eziyise?	◯ Yes ◯ No
Have you carried more than one passenger per trip in the past 30 days?/Wali oweeseko ku basabaze abassukka mu omu munnaku 30 eziyise?	◯ Yes ◯ No

Outcome Information

(all questions refer to while you are working as a boda-boda driver/Bino ebibuzo byonna bikwatagana ku banga/obudde nga okola nga omuvuzi wa boda boda)

Do you ever wear a helmet while driving a motorcycle?/Oyambala elementi nga ovuga piki piki?	◯ Yes ◯ No
In the past 30 days, how often would you say you wore a helmet while driving a motorcycle?/Munnaku asaatu eziyise, olowoza Oyambaliddemu elementi emirundi emmeka?	◯ Never/Tewali mulundi nagumu ◯ Sometimes/Emirundi egiimu ◯ Always (all the time)/Emirundi gyonna ◯ Don't know (unsure)/Simanyi ◯ Refused to answer/Sijja kuddamu ◯ Do not own a helmet/Sirina elementi

Exposure Information (to be completed by the interviewer/byakujjuzibwaamu oyo abuuza)

Is the subject wearing anything to identify him as a SafeBoda driver?/Eyetabye mukunonyereza alika ekintu kyonna kya'yamadde ekiraga nti muvuzi wa Safe boda?

◯ Yes
◯ No

If Yes, which one? (mark all that apply)/Bwekibanga kituufu, kinyonyole?

☐ Sticker/Sitiika
☐ Jacket/Jakeeti
☐ Other/Ekirara kyonna

Exposure Information Questions for the Participant/Byakujjuzibwa abuuziibwa

Were you ever a SafeBoda driver?/Wali abaddeko mu kibiina kta Safe Boda?

◯ Yes
◯ No

Are you currently a driver for the SafeBoda program?/Oli muvuzi wa Safe boda kati?

◯ Yes
◯ No

When did you become a SafeBoda driver?/Wafuuka ddi omuvuzi wa Safe Boda?

____

Do you currently use the SafeBoda app for accepting passenger request for rides?/Mukiseera kino okozesa kunkola ya Safeboda ey'okumutimbagano okufuna abasabaze?

◯ Yes
◯ No

How often would you say you wore a helmet while a motorcycle before you became a SafeBoda driver?/Mirundi gyenkanaki gyo'yinza okugamba nti wayambala nga elementi nga tonafuuka muvuzi wa Safeboda?

◯ Never/Tewali mulundi nagumu driving
◯ Sometimes/Emirundi egiimu
◯ Always (all the time)/Emirundi byonna
◯ Don't know (unsure)/Simanyi
◯ Refused to answer/Sijja kuddamu
◯ Did not own a helmet/Sirina elementi

Do you have a friend or a peer who is a SafeBoda driver?/Olinayo mukwano gwo yenna omuvuzi wa Safeboda?

◯ Yes
◯ No

Has having a friend or peer who is a SafeBoda driver impacted your use of any of the following? (mark all that apply)/Okuba ne mukwano gwo omuvuzi wa Safeboda, kikoseza kitya enkozessayo eyabino wammanga?

☐ Helmet/Elementi
☐ Reflective jacket/Jaketi
☐ Not at all/Tewali yadde

Cohort study of RTC and RTI Incidence Follow-up Information/Ebinayamba okwongera okunoonyereza

Primary mobile phone number/Ennamba ye'ssimu esooka

____

Secondary mobile phone number/Ennamba ye'ssimu ey'okubiri

____

Date of cohort enrollment/Ennaku zo'mwezi omuntu walondebbwa

____
(Enter date of cohort enrollment here/Wandiika wano ennaku zo'mwezi)

Riding Behavior Study Consent Information

(Please complete this section after receipt of verbal consent from the participant for the cohort study)/Jjuza bino oluvanyuma lwo;mubuziibwa okukiriza

Participant consented to participate in pilot study to measure driving patterns among boda-boda drivers in Kampala/Eyetabye mukunonyereza yakkiriza okwetabamu okusobola okulaba engeri abavuzi ba boda boda ab'omu Kampala gyebavugamu?	◯ Yes ◯ No
Date of GPS tracker and accelerometer installation/Ennaku zo'mwezi obuuma bu kalimageezi lwebunaatekebwako	__________________________ (Can be scheduled in advance/Esobola okuteesebwaako nga bukyaali)
GPS tracker ID/Ennamba ya'kuuma akalaga ekifo	__________________________
Accelerometer ID/Ennamba ya'kuuma akapiima omuliiro	__________________________

Figure S2 Road traffic knowledge assessment tool (adapted from the Uganda Highway Code)

Cohort Traffic Knowledge Questionnaire

Record ID	______________________
Participant ID/Ebikwaata kwoyo abuuzibwa	______________________ (Must enter the same ID as in the helmet use questionnaire/Erina okufanana ne'yayingizibwa mu bibuuzo ebikwata kunkozesa ya elementi)
Language questionnaire administered in/Olulimi olukozesebbwa	◯ English ◯ Luganda

---

Road Traffic Knowledge Questions (from the Uganda Highway Code)/Ebibuuzo ku bumanyi bwa'mateeka g'okulugudo. (Okuva mu Uganda Highway Code)

Interviewer: Read each statement loud to the participant including the three response options. DO NOT prompt or suggest a response to the participant/Ssomera abuziibwa ebibuuzo wamu ne'byokuddamu kwalina okulonda. Togezaako okumubulira oba okumusendasenda.

Uganda has a helmet law that mandates helmet use riding a motorcycle/Uganda erina etteka erikakata kunkozesa ya elementi ngo'vuga piki piki	◯ True/Kituufu while ◯ False/Kikyaamu ◯ Don't Know/Simanyi
Overtaking of other vehicles is illegal unless you clear and unobstructed view of the road ahead/Okuyisa emmottoka endala kuluguudo kimenyamateeka okujjako ng'olaba bulungi nnyo gy'olaga	◯ True/Kituufu have a ◯ False/Kikyaamu ◯ Don't Know/Simanyi
It is okay to wear sandals or slippers when driving a motorcycle/Tekirina buzibu okwambala ssilipa oba engatto eziraga ebigere ng'ovuga piki piki.	◯ True/Kituufu ◯ False/Kikyaamu ◯ Don't know/Simanyi
Every boda-boda driver must reduce speed upon approaching any crossroads or intersection/Buli muvuzi wa boda boda alina okukendeza sipiidi bw'atuuka mu massanganzira.	◯ True/Kituufu ◯ False/Kikyaamu ◯ Don't Know/Simanyi
A boda-boda driver does not have to stop for a pedestrian who has already stepped on a zebra crossing/Tekyetaagisa muvuzi wa boda boda kuyimirira kufuna musabaze abamaze okulinya webasalira oluguudo.	◯ True/Kituufu ◯ False/Kikyaamu ◯ Don't Know/Simanyi
The maximum speed limit for a boda-boda driving through an urban or trading center in Uganda is 50km per hour/Buli muvuzi wa boda mu Uganda talina kusussa km 50 buli ssawa bwaaba avuga mu kibuga oba mu kabuga.	◯ True/Kituufu ◯ False/Kikyaamu ◯ Don't Know/Simanyi
The yellow/amber light in a traffic signal means "stop unless you are so close to the stop line that it would be dangerous to try and stop"/Ettala lya kyenvu elyokulugudo litegeeza " Yimirira okuleka nga oli kumpi nnyo ne layiina era nga kijja kuba kyaabulabe bwoyimirira.	◯ True/Kituufu ◯ False/Kikyaamu ◯ Don't Know/Simanyi

It is not required for a boda-boda driver to have a valid driving permit to operate a motorcycle in Uganda/Tekyetaagisa muvuzi wa boda kuba nakiwaandiiko ekimukkiriza okuvuga piki piki mu Uganda.

- ○ True/Kituufu
- ○ False/Kikyaamu
- ○ Don't Know/Simanyi

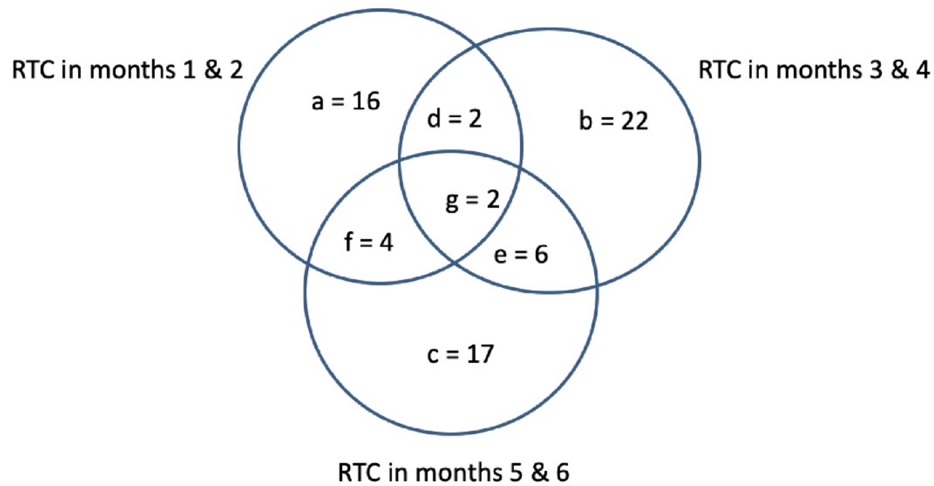
Every motorcycle must be insured and registered in Uganda/Buli piki piki erina okuba ne yinsuuwa ate era nga mpaandiise mu Uganda.

- ○ True/Kituufu
- ○ False/Kikyaamu
- ○ Don't Know/Simanyi

A boda-boda driver must keep at least one hand on the handlebars except when signaling/Buli muvuzi wa boda alina okukuuma wakiri omukono gumu ku piki piki okugyako nga'lina obubaka bwa'wereeza.

- ○ True/Kituufu
- ○ False/Kikyaamu
- ○ Don't Know/Simanyi

Figure S3 Venn diagram showing distribution of road traffic crashes during the six months of follow-up



Legend

a= # RTC in months 1 & 2 only (first 2-month interval)
b= # RTC in months 3 & 4 only (second 2-month interval)
c= # RTC in months 5 & 6 only (third 2-month interval)
d= # RTC in both months 1 & 2 and 3 & 4
e= # RTC in both months 3 & 4 and 5 & 6
f= # RTC in both months 1 & 2 and 5 & 6
g= # RTC in each 2-month interval of follow-up

Figure S1 Computer-assisted personal interview (CAPI) questionnaire

Study: Motorcycle taxis and road safety in Kampala: comparing health outcomes in SafeBoda and regular motorcycle taxi drivers

Ebibuuzo kunkozesa ya elementi

Record ID ______________________

Interviewer and Study Site Information/Ebikwata ku abuuza ne'kifo

Initials of interviewer/Ammanya g'omunonyereza mu bufunze ______________________

Location of boda-boda stage/Ekifo siteegi ya boda we'sangibwa ______________________
(Describe where the stage is located (include cross streets)/Nyonyola ekifo, ekitundu ne' nguudo eziriranyewo)

Boda-boda stage ID/Endaga kifo ya siteegi ______________________

Name of the part of the city where the stage is located/Erinnya lye'kitundu mu kibuga siteegi weesangibwa ______________________

Number of boda-boda drivers at the stage/Omuwendo gwa'bagoba ba boda ku siteegi ______________________

Number of boda-boda drivers at the stage with visible helmets (either worn or strapped on their motorcycle)/Omuwendo gwa'bagoba/abavuzi ba boda abalina elementi za piki ezirabika (abazambade oba abatazambade ngo'zilabako) ______________________

Other Information (to be completed right before the interview)/Ebintu ebirara ebilina okudibwamu nga omunonyereza tanatandika kunonyereza/kubuliriza

Participant ID/Namba yeyetabye mukunonyereza ______________________

Language questionnaire administered in/Olulimi olubuziibwamu ebibuuzo
- ◯ English
- ◯ Luganda

Date of interview/Ennakku zo'mweezi ______________________

Demographic and Other Covariate Information/Ebikwata kumbeera yabantu bomukitundu

Sex of the participant/Ekikula ky'omuntu eyetabyeemu
- ◯ Female/Mukazi
- ◯ Male/Musajja

How old are you?/Olina emyaaka emeka? ______________________
(Age in years)

How many years of schooling have you completed? (exclude repeat years)/Wasoma kwenkana ki? ______________________

What is your marital status?/Embeera yo ekwatagana nebyobufumbo eyimiridde etya?

- ◯ Single, never married/Towasanga
- ◯ Married or co-habiting/Mufumbo a mubeera mwena
- ◯ Widowed/Ssemwandu
- ◯ Separated/Mwayawukana
- ◯ Divorced/Mwayawukana mu mateeka

Do you have other job (s) apart from working as a boda-boda driver?/Olinayo omulimu/emiriimu emilara okuleka ogwo'buvuzi bwa boda boda?

- ◯ Yes
- ◯ No

What other job (s) do you currently have?/Mukiseera kino milimu ki emirara gy'olina?

____

What is your gross average weekly income as a boda-boda driver?/Okutwaliza awamu nga omuvuzi wa boda boda, ofuna sente meka buli sabiiti/wiiki?

____

What is your net average weekly income as a boda-boda driver (that is after subtracting any expenses)?/Okutwaliza awamu, ngomuvuzi wa boda boda ofuna sente meka buli sabiiti/wiiki ngo'jjeko empooza, ne'bilara?

____

What hours do you usually work as a boda-boda?/Otera kukola budde ki obw'olunaku?

- ◯ Day hours only (6am to 6pm)/Missana
- ◯ Night hours only (6pm to 6am)/Kiro
- ◯ Both day and night hours/Missana na Kiro

How many hours per day do you on average work as a boda-boda driver?/Mukugerageranya mulunaku okola essawa mmeka nga omuvuzi wa boda boda?

____
(Enter in hours (no decimals)/Wandiika essawa mubujuvu)

How many trips, on average, do you make per day carrying passengers?/Ovuga luutu z'abasabaze meka mulunaku?

____

Do you drink alcohol?
(by alcohol we mean beer, wine, or liquor in a sachet or bottle)/Onnywa omwenge?

- ◯ Yes
- ◯ No

Have you driven your motorcycle within 2 hours of drinking alcohol in the past 30 days?/Munaku asaatu eziyise, wali ovuze ku piki mu saawa bbiri oluvanyuma lw'okunywa omwenge?

- ◯ Yes
- ◯ No

What type of motorcycle do you have?/E piki piki gyo'lina yakika ki?

- ◯ Bajaj
- ◯ Boxer
- ◯ Senke
- ◯ Yamaha
- ◯ Honda
- ◯ Other

What's the engine size of your motorcycle/Yingini ya piki piki yo yaamanyi ki?

- ◯ Less than 100cc/Wansi wa cc 100
- ◯ 100cc or greater/cc 100 no'kwambuka

Have you used your mobile phone while actively driving your motorcycle in the past 30 days?
(excludes use while stopped)/Wali okozeseza ku simu okukuba oba okuwereeza obukaka nga bwovuga munnaku asatu eziyise?

- ◯ Yes
- ◯ No

Do you currently own a helmet for your motorcycle?/Olina elementi ey'obwannanyini?

- ◯ Yes
- ◯ No

What's the type of your helmet?/Elementi yo ya kikaa ki?

____
(Write the brand name and then the type in parenthesis/Wandiika omukozi wa elementi)

What's the color of your helmet?/Elementi yo yaalangi ki?

____

How long have you had your helmet?/Elementi yo omaze nayo banga ki?

____

How much (in UGX) did you pay for your helmet? (applicable only to those who purchased their helmet, not given for free)/Elementi yo wagigula ssente mmeka eza Uganda?

____

Do you currently own any of these protective gear to wear while driving your motorcycle? (mark all that apply)/Olina yo kubyokwekumisa bino wammanga byoosobola okwambala ngo'vuga piki piki?

- ☐ Sturdy boots/Engatto ezivuga
- ☐ Gloves/Gilaavuzi
- ☐ Reflective jacket/Ekkoti eyaka/jaketi
- ☐ Helmet/Elementi
- ☐ Protection glasses/Ggalubindi
- ☐ Thick pants/trousers/Empale engumu
- ☐ None/Tewali

Do you own the motorcycle you use for boda-boda?/Piki piki gyovuga yiyo?

- ◯ Yes
- ◯ No

Are you a member of any of these boda-boda programs? (mark all that apply)/Oliko mu nteka teka/pulogulaamu yonna eyabavuzi ba boda boda?

- ☐ SafeBoda
- ☐ BodaPro
- ☐ Tugende
- ☐ None/Tewali

Have you ever received a road safety training from programs such as Uganda Helmet Vaccine Initiative, SafeBoda, BodaPro, or Tugende?/Wali otendekeddwako kunkozesa yolugudo okuva mu bibiina nga Uganda Helment Vaccine initiative, Safeboda, Bodapro oba Tugende?

- ◯ Yes
- ◯ No

How long (in years) have you worked as a boda-boda driver?/Omulimu gwo'kuvuga boda ogumazeemu banga ki?

____

Are you a member of any boda-boda union or association?/Olina ekibiina kyona ekyabavuzi ba boda boda kyolimu?

- ◯ Yes
- ◯ No

What's the name of your union or association?/Ekibiina bakiyita batya?

____

Do you have a driver's license?/Olina ebisanyizo ebikukiriza okuvuga boda?

- ◯ Yes
- ◯ No

How did you learn how to drive a motorcycle?/Wayiga otya okuvuga pikipiki?

- ◯ Self taught (received no help from other people)/Weka
- ◯ Friend/relative/peer/Mukwano/ owo'luganda
- ◯ Driving school/Muttendekero lya'bavuzi
- ◯ Other/Ewalala wonna

At what age did you start driving a motorcycle?/Watandika okuvuga piki ngo'lina emyaka emeka?

____

How long (in days) did it take you from when you started driving a motorcycle to when you started driving it on public roads?/Kyakutwalira banga ki okuva bwewatandiika okuyiga okutuusa lwewatandika okuvugira ku nguudo ezikozesebwa abantu?

____

How long (in weeks) did it take you from when you started driving a motorcycle to when you started carrying passengers as a boda-boda driver?/Kyakutwalira ssabiiti mmeka okuva lwewatandiika okuyiga paka bwe watandiika okutwala abantu nga omuvuzi wa boda boda?

____

Were you involved in a road traffic crash in the past 6 months (by crash we mean an incident where you collided with another vehicle (e.g. car, motorcycle), a pedestrian, an animal, or an object)/Wali ofunye ku kabenje akekika kyonna mumyeezi 6 egiyise? (akabenje tutegeeza wewekoona nekidukka ekirala, nga emotoka,pikipki, omusabaze,ensolo oba ekintu kyonna).

◯ Yes
◯ No

Who were the other parties involved in your most recent crash? (mark all that apply)/B'ani abalala abali mu kabenje ako?

☐ Motorcycle/Piki piki
☐ Car/truck/Mmotoka
☐ Bus/Baasi
☐ Bicycle/Akagaali
☐ Pedestrian/Omutanbuuze
☐ Animal/Ensolo
☐ Tree/pole/road barriers/Omuuti/Ekikondo
☐ None/Tewali

Were you wearing a helmet at the time of your most recent crash?/Wali oyanbade elementi mu kabenje kewakasembayo okufuna?

◯ Yes
◯ No

At the time of your most recent crash, which of the following was true of you? (mark all that apply)/Mukiseera kyewafuniramu akabenje, kiki ku bino Kyandiiba nga kye'kyakuleetera akabenje?

☐ Speeding/Ndiima
☐ Faulty brakes/Ebiziyiza ebifu
☐ Making an illegal turn/Okukyukira awakyamu
☐ Slippery surface/Obuseerezi
☐ Drink driving/Obutamiivu
☐ Running a red light or stop sign/Obutagondera bitaala
☐ Wrong way driving/Okuvugira ku ludda olukyaamu
☐ Distracted driving/Obutassayo Mwoyo
☐ Other/Ekirara kyonna

Who do you think was at fault for your most recent crash?/Ani gwo'lowooza eyali mu nsobi mu kabenje kewasembayo okufuna?

◯ Myself/Nze mwennyini
◯ The other party (ies) involved/Abalala abali mukabenje
◯ No one/Tewali
◯ Not sure/Simanyi bulungi

Where did your most recent crash occur?/Akabenje kewasembayo okufuna wakafunira luddawa?

____
(Road or street name where the crash happened (ask for cross streets or a landmark if possible)/Erinnya lyo'lugudo awaagwa akabenje))

When did your most recent crash occur?/Wafuna ddi akabenje ako?

◯ Morning (6am-12pm)/Kumakya
◯ Afternoon (12pm-6pm)/Mu ttuntu
◯ Night (6pm-6am)/Ekiro

Did you sustain an injury from your most recent crash?/Wafuna mu ebisago okuva ku kabenje kewakasembayo okufuna?

◯ Yes
◯ No

What body part (s) did you injure? apply)/Bitundu ki kumubiri ebyakosebwa?	☐ Head/Mutwe (mark all that ☐ Neck/Nsiingo ☐ Arms/hands/Mikono/engalo ☐ Chest/Kifuba ☐ Legs/Magulu ☐ Pelvis/Enkoggo ☐ Feet/Bigere ☐ Spinal cord/Enkiizi ☐ Other/Ekirara kyonna
Did your injury require a visit to a health facility to get care?/Oluvanyuma lwa'kabenje, kyali kikwetagisa okufuna obujanjabi okuva mu ddwaliro?	◯ Yes ◯ No
Did your injury require an admission (that is inpatient care)?/Obuvune bwo bwali bwetagisa okuweebwa ekitanda?	◯ Yes ◯ No
For how many days were you admitted at the healthcare facility for your injury?/Wamala ennaku meka ku kitanda?	______________________
Have you driven your motorcycle on a curb or pavement in the past 30 days?/Wali ovugiddeko piki piki yo ku peevumenti munnaku 30 eziyise?	◯ Yes ◯ No
Have you driven your motorcycle in the wrong direction (opposite to oncoming traffic) in the past 30 days?/Wali ovugiddeko piki piki yo ku ludda olukyamu munnaku 30 eziyise?	◯ Yes ◯ No
Have you carried more than one passenger per trip in the past 30 days?/Wali oweeseko ku basabaze abassukka mu omu munnaku 30 eziyise?	◯ Yes ◯ No

Outcome Information

(all questions refer to while you are working as a boda-boda driver/Bino ebibuzo byonna bikwatagana ku banga/obudde nga okola nga omuvuzi wa boda boda)

Do you ever wear a helmet while driving a motorcycle?/Otera okwambala elementi nga ovuga piki piki?	◯ Yes ◯ No
In the past 30 days, how often would you say you wore a helmet while driving a motorcycle?/Munnaku asaatu eziyise, oyambaliddeme elementi emirundi emmeka?	◯ Never/Tewali mulundi nagumu ◯ Sometimes/Emirundi egiimu ◯ Always (all the time)/Emirundi gyonna ◯ Don't know (unsure)/Simanyi ◯ Refused to answer/Sijja kuddamu ◯ Do not own a helmet/Sirina elementi

Exposure Information (to be completed by the interviewer)/byakujjuzibwaamu oyo abuuza

Is the subject wearing anything to identify him as a SafeBoda driver?/Eyetabye mukunonyereza alika ekintu kyonna kya'yamadde ekiraga nti muvuzi wa Safe boda?

◯ Yes
◯ No

If Yes, which one? (mark all that apply)/Bwekibanga kituufu, kinyonyole?

☐ Sticker/Sitiika
☐ Jacket/Jaketi
☐ Other/Ekirara kyonna

Exposure Information Questions for the Participant/Byakujjuzibwa abuuziibwa

Were you ever a SafeBoda driver?/Wali abaddeko mu kibiina kta Safe Boda?

◯ Yes
◯ No

Are you currently a driver for the SafeBoda program?/Oli muvuzi wa Safeboda kati?

◯ Yes
◯ No

When did you become a SafeBoda driver?/Wafuuka ddi omuvuzi wa Safeboda?

____

Do you currently use the SafeBoda app for accepting passenger request for rides?/Mukiseera kino okozesa kunkola ya Safeboda ey'okumutimbagano okufuna abasabaze?

◯ Yes
◯ No

How often would you say you wore a helmet while a motorcycle before you became a SafeBoda driver?/Mirundi gyenkanaki gyo'yinza okugamba nti wayambala nga elementi nga tonafuuka muvuzi wa Safeboda?

◯ Never/Tewali mulundi nagumu driving
◯ Sometimes/Emirundi egiimu
◯ Always (all the time)/Emirund byonna
◯ Don't know (unsure)/Simanyi
◯ Refused to answer/Sijja kuddamu
◯ Did not own a helmet/Sirina elementi

Do you have a friend or a peer who is a SafeBoda driver?/Olinayo mukwano gwo yenna omuvuzi wa Safeboda?

◯ Yes
◯ No

Has having a friend or peer who is a SafeBoda driver impacted your use of any of the following? (mark all that apply)/Okuba ne mukwano gwo omuvuzi wa Safeboda, kikoseza kitya enkozessayo eyabino wammanga?

☐ Helmet/Elementi
☐ Reflective jacket/Jaketi
☐ Not at all/Tewali yadde

Cohort Study of RTC and RTI Incidence Consent Information

(Please read the cohort study consent form before completing this section)

Participant consented to participate in cohort study/Eyetabyeemu akkiriza okwetaba mu musomo

◯ Yes
◯ No

Primary mobile phone number/Ennamba ye'ssimu esooka

____

Secondary mobile phone number/Ennamba ye'ssimu eyokubiri

____

Date of cohort enrollment/Ennaku zo'mwezi omuntu mwa'londebbwa

____
(Enter date of cohort enrollment here/Yingiiza ennaku zo'mwezi wano)

Riding Behavior Study Consent Information

(Please complete this section after receipt of verbal consent from the participant for the cohort study/Juuza bino oluvanyuma lwo'mubuziibwa okukkiriza okwetaba mu kusoma kuno).

Participant consented to participate in pilot study to measure driving patterns among boda-boda drivers Kampala/Omubuziibwa akkiriza okwetaba mu kusoma kunvuga ya boda mu bavuzi be'Kampala.	◯ Yes ◯ No in
Date of GPS tracker and accelerometer installation/Ennaku zo'mwezi obuuma bu kalimageezi lwebunaatekebwako	__________________________ (Can be scheduled in advance/Esobola okuteesebwako nga bukyaali)
GPS tracker ID/Enyinyonyola ya'kuuma akalaga ekifo	__________________________
Accelerometer ID/Enyinyonyola ya'kuuma akapiima omuliiro	__________________________

Figure S2 Sample R code from processing and analysis of the GPS data

```
# setups
library(here, verbose = FALSE)
library(RPostgreSQL)
library(sqldf)
library(tcltk)
library(lubridate)
options(stringsAsFactors = FALSE)

# pull parameters from the params.R file
projectdir <- here::here()
scriptsdir <- file.path(projectdir, "tools")
datadir <- file.path(projectdir, "data")
source(file.path(projectdir, "params.R"))
password <- scan("~/.pgpasswd", what = "character", quiet = TRUE)

# db --------------------------------------------------------------------
# db connection function
# assumes we are only using one database for the project, with dbname = projectname and
connection = "dbconn"
connectDB <- function(dbname=mydbname, host="localhost", port=5433){
  # read the password file
  if(!exists("dbconn")){
    #message(dbname, " not connected")
    dbconn <- dbConnect(dbDriver("PostgreSQL"), dbname = mydbname, host=host, password
= password, port=port)
    message(dbname, " now connected as dbconn")
    return(dbconn)
  }
  if(exists("dbconn")){
    connectionIsCurrent <- isPostgresqlIdCurrent(dbconn)
    #message(connectionIsCurrent)
    if(!connectionIsCurrent){
      message("establishing connection to ", dbname)
      dbconn <- dbConnect(dbDriver("PostgreSQL"), dbname = mydbname, host=host,
password = password, port=port)
      #message(dbname, " now connected")
    }
    return(dbconn)
  }
}
```

```
# db connect ----------------------------------------------------------------
# connect to the db
dbconn <- boda <- connectDB(dbname="boda")

# sqldf options for Postgres
mypassword <- scan("~/.pgpasswd", what = "character", quiet = TRUE)
options(sqldf.RPostgreSQL.user = Sys.getenv("username"),
        sqldf.RPostgreSQL.password = password,
        sqldf.RPostgreSQL.dbname = mydbname,
        sqldf.RPostgreSQL.hostname = "localhost",
        sqldf.RPostgreSQL.port = 5432)


mvExists <- function(conn, schemaname, mvname){
    sql <- sprintf(
"SELECT count(*) > 0 as mvexists
FROM pg_catalog.pg_class c
     LEFT JOIN pg_catalog.pg_namespace n ON n.oid = c.relnamespace
WHERE c.relkind IN ('v','s','')
      AND n.nspname !~ '^pg_toast'
  AND c.relname ~ '^%s$'
  AND n.nspname ~ '^%s$';", mvname, schemaname)
    dbGetQuery(conn = conn, statement = sql)$mvexists
}

tExists <- function(conn, table_schema, table_name){
    sql <- sprintf("select count(*) = 1 from information_schema.tables where table_schema = '%s'
and table_name = '%s';", table_schema, table_name)
    dbGetQuery(conn = conn, sql)[1,1]
}


# read in the GPS data

library(lubridate)
library(here)
library(sqldf)
library(RPostgreSQL)
library(RSQLite)
library(sf)

projectdir <- here::here()
source(file.path(projectdir, "tools", "setup.R"))
```

```
gps.dir <- file.path(projectdir, "data", "gps")
gpsfiles <- sort(list.files(gps.dir, "*.csv$", full.names = TRUE, recursive = TRUE))
files_ids <- data.frame(fname = gpsfiles, id = gsub(".csv", "", tolower(do.call(c,
lapply(strsplit(gpsfiles, split = "_"), function(x) x[2])))))
files_ids <- files_ids[order(files_ids$id),]


maketablestructure_gps <- function(conn = dbconn, schema){
  sql <- "
  DROP TABLE if exists xxxSCHEMAxxx.gps cascade;
  CREATE TABLE xxxSCHEMAxxx.gps
  (
    id text,
    wave text,
    index integer,
    rcr text,
    utc_date text,
    utc_time text,
    ms integer,
    valid text,
    height_m double precision,
    speed_kmh double precision,
    heading double precision,
    dsta integer,
    dage integer,
    pdop double precision,
    hdop double precision,
    vdop double precision,
    nsat_used text,
    sat_info text,
    distance_m text,
    nsat_view text,
    time_gps timestamp with time zone,
    time_gps_std timestamp with time zone,
    jdaya_gps date,
    gps_src text,
    latitude double precision,
    longitude double precision,
    the_geom_4326 geometry(POINTM, 4326),
    the_geom_xxxSRIDxxx geometry(POINTM, xxxSRIDxxx),
    gid serial
  );
  alter table xxxSCHEMAxxx.gps add primary key (gid);
  create index idx_gps_id_wave on xxxSCHEMAxxx.gps using btree(id, wave);
```

```
   create index idx_gps_id on xxxSCHEMAxxx.gps using btree(id);
   create index idx_gps_time_gps_std on xxxSCHEMAxxx.gps using btree(time_gps_std);
   create index idx_gps_the_geom_4326 on xxxSCHEMAxxx.gps using gist(the_geom_4326);
   create index idx_gps_the_geom_xxxSRIDxxx on xxxSCHEMAxxx.gps using
gist(the_geom_xxxSRIDxxx);
   create index idx_gps_src on xxxSCHEMAxxx.gps using btree(gps_src);
   create index idx_gps_id_wave_time on xxxSCHEMAxxx.gps using btree(id, wave,
time_gps_std);
   "

   # substitute for the correct SRID
   sqlx <- gsub("xxxSCHEMAxxx", schema, gsub("xxxSRIDxxx", mysrid, sql))
   O <- dbGetQuery(conn = conn, statement = sqlx)
}


make_child_tables <- function(schema){
   for(i in 1:nrow(files_ids)){
      fname <- files_ids$fname[i]
      id <- files_ids$id[i]
      message(i, " of ", length(gpsfiles), " (", basename(fname), ")")
      sqli <- sprintf("create table %s.gps_%s (
                  check (id = '%s'::text)
      ) inherits (%s.gps);", schema, id, id, schema)
      O <- dbGetQuery(conn = dbconn, statement = sqli)
      O <- dbGetQuery(conn = dbconn, statement = sprintf("create index idx_gps_%s on
%s.gps_%s using btree(id);", id, schema, id))
      O <- dbGetQuery(conn = dbconn, statement = sprintf("create index idx_gps_%s_geom4326
on %s.gps_%s using gist(the_geom_4326);", id, schema, id))
      O <- dbGetQuery(conn = dbconn, statement = sprintf("create index
idx_gps_%s_geom32636 on %s.gps_%s using gist(the_geom_32636);", id, schema, id))
   }

   tsql0 <- "
drop function gps_insert_trigger;
CREATE OR REPLACE FUNCTION gps_insert_trigger()
RETURNS TRIGGER AS $$
BEGIN"

   for(i in 1:nrow(files_ids)){
      id <- files_ids$id[i]
      if(i==1){
         tsql0 <- sprintf("%s\n
         IF ( NEW.id = '%s' ) THEN
```

```
      INSERT INTO %s.gps_%s VALUES (NEW.*);", tsql0, id, schema, id)
    } else {
      tsql0 <- sprintf("%s\n
      ELSIF ( NEW.id = '%s' ) THEN
      INSERT INTO %s.gps_%s VALUES (NEW.*);", tsql0, id, schema, id)
    }

  }

  tsql0 <- sprintf("%s\n
    END IF;
    RETURN NULL;
END;
$$
LANGUAGE plpgsql;", tsql0)

  cat(tsql0, file = "/tmp/tf.sql")

}
# read in the GPS file
qstarz.to.db.afile <- function(fname, conn, schema, overwrite = FALSE){
  violation <- FALSE
  message(sprintf("   processing in %s ....", fname))
  # id and wave number, get from params.R
  id <- tolower(unlist(strsplit(basename(fname), split="[[:punct:]]"))[fnchunk.nameparts$id])
  wavenum <- unlist(strsplit(basename(fname), split="[[:punct:]]"))[fnchunk.nameparts$wave]

  # wave
  wave <- ifelse(wavenum==1, "baseline",
    ifelse(wavenum==2, "post1",
        ifelse(wavenum==3, "post2", NA)))

  # do records for this exist?
  hasrecs <- dbGetQuery(conn = conn, statement = sprintf("select count(*) > 1 as hasrecs from
%s.gps where gps_src = '%s';", schema, fname))$hasrecs

  # if records exist
  if(hasrecs){
    # if we don't want to overwrite
    if(!overwrite){
      message("   not overwriting data for ", fname)
      return(invisible())
    } else {
```

```
        O <- dbGetQuery(conn = conn, statement = sprintf("delete from %s.gps where gps_src = 
'%s';", schema, fname))
      }
   }

   # zero size?
   if(file.info(fname)$size == 0){
      message("   BAILING! ", fname, " is size zero.")
      return(invisible())
   }
   # read data
   gps <- read.csv(fname, as.is=T, strip.white=T)
   # if there are fewer than 100 records then bail
   if(nrow(gps) < 100){
      message("   BAILING! only ", nrow(gps), " records in ", fname)
      return(invisible())
   }
   # column names
   colnames(gps) <- gsub("_$", "", tolower(gsub("[[:punct:]]", "_", colnames(gps))))
   # sat info column
   colnames(gps)[grep("sat_info", colnames(gps))] <- "sat_info"
   # drop null records
   gps <- gps[grep("-", gps$utc_date, invert=T),]
   # reformat lat & long
   gps$latitude <- ifelse(gps$n_s=="S", -gps$latitude, gps$latitude)
   gps$longitude <- ifelse(gps$e_w=="W", 0 - gps$longitude, gps$longitude)
   # some points are not valid
   gps <- gps[gps$valid != "NO FIX",]
   # add the subject ID
   gps <- data.frame(id, wave, gps)
   # does this meet specification?
   if(any(grepl("altitude", names(gps)))){
      violation <- TRUE
      colnames(gps)[grep("altitude", colnames(gps))] <- "height"
      colnames(gps)[grep("track_id", colnames(gps))] <- "rcr"
      gps$nsat_used <- "NA/NA"
      # drop unnecessar columns
      gps$g_x <- gps$g_y <- gps$g_z <- NULL
      # add "placeholder" columns
      gps$dsta <- gps$dage <- gps$pdop <-gps $vdop <-gps $hdop <- gps$distance_m <- NA
   }
   # reformat height, speed, distance, nsat used, nsat view: remove units from values and add 
to field names
   gps$height <- as.numeric(sub("M", "", gps$height))
```

```
colnames(gps)[grep("height", colnames(gps))] <- "height_m"
gps$speed <- as.numeric(sub("km/h", "", gps$speed))
colnames(gps)[grep("speed", colnames(gps))] <- "speed_kmh"
# sometimes distance was strangely named and formatted.
if(any(grepl("^distance$", colnames(gps)))){
   gps$distance <- sub(" M", "", gps$distance)
   colnames(gps)[grep("^distance$", colnames(gps))] <- "distance_m"
}
# split nsat into view and used
colnames(gps)[grep("nsat", colnames(gps))] <- "nsat_used"
gps$nsat_used <- gsub(" ", "", gps$nsat_used)
gps$nsat_view <- gps$nsat_used
gps$nsat_used <- unlist(strsplit(gps$nsat_used, "\\/"))[1]
gps$nsat_view <- unlist(strsplit(gps$nsat_view, "\\/"))[2]
if(violation){
   gps$nsat_used <- gps$nsat_view <- NA
}
# handle timestamps: create a single UTC full timestamp
gps$time_gps <- with_tz(as.POSIXct(sprintf("%s %s", gps$utc_date, gps$utc_time),
format="%Y/%m/%d %H:%M:%S", tz="UTC"), "PST8PDT")
# round time to 1 minute
gps$time_gps_std <- floor_date(gps$time_gps, unit = "1 min")
# date, transition at 3 AM
gps$jdaya_gps <- as.Date(gps$time_gps - 3 * 60 * 60, tz = timezone)

# drop records from 1980
gps <- gps[format(gps$time_gps, "%Y")!="1980",]
# source
gps$gps_src <- fname

# drop junk columns
gps$local_date <- gps$local_time <- gps$n_s <- gps$e_w <- NULL

# drop trailing underscore on column names
colnames(gps) <- gsub("_$", "", colnames(gps))
gps$secs <- as.numeric(gps$time_gps)

gps_sf = st_as_sf(gps, coords = c("longitude", "latitude", "secs"), dim = "XYM", crs = 4326, agr
= "constant")
gps_sf$latitude <- st_coordinates(gps_sf$geometry)[,"Y"]
gps_sf$longitude <- st_coordinates(gps_sf$geometry)[,"X"]
# wkb columns for writing to PostGIS
# WGS84
gps_sf$the_geom_4326 <- st_as_binary(gps_sf$geometry, hex = TRUE, EWKB = TRUE)
```

```
    # projected
    gps_sf$the_geom_proj <- st_as_binary(st_transform(gps_sf$geometry, mysrid), hex = TRUE,
EWKB = TRUE)
    colnames(gps_sf)[grep("the_geom_proj", colnames(gps_sf))] <- sprintf("the_geom_%s",
mysrid)
    gps_sf$geometry <- NULL
    gps$gid <- NULL

    assign("gps", gps, envir = .GlobalEnv)

    #

    # write to database
#    O <- dbWriteTable(conn = conn, name = c("gps", "gps"), gps_sf, row.names = FALSE,
overwrite = FALSE, append = TRUE, field.types = NULL, factorsAsCharacter = TRUE, binary =
TRUE)
    O <- dbWriteTable(conn = conn, name = c(schema, "gps"), gps_sf, row.names = FALSE,
overwrite = FALSE, append = TRUE, field.types = NULL, factorsAsCharacter = TRUE, binary =
TRUE)

    # return a data frame of the GPS
    return(gps_sf)
}

# push many GPS CSV files to the database
qstarz.to.db.allfiles <- function(conn = dbconn, gpsfiles = gpsfiles, schema, overwrite = FALSE,
where=""){   #WHERE valid='DGPS' AND hdop < 5", verbose=F){
    if(overwrite){
        maketablestructure_gps(conn = conn)
    }
    for(i in 1:length(gpsfiles)){
        fname <- gpsfiles[i]
        message(i, " of ", length(gpsfiles), " (", basename(fname), ")")
        x <- qstarz.to.db.afile(fname = fname, conn = conn, schema = schema, overwrite =
overwrite)
    }
}

# make GPS segments between points
gps.segments <- function(conn = dbconn){
    sql <- "
    drop table if exists xxxSCHEMAxxx.gps_segments;
    create table xxxSCHEMAxxx.gps_segments as
```

```
   with g as (select *, row_number() over(order by wave, id, time_gps) as xid from
xxxSCHEMAxxx.gps order by wave, id, time_gps)
   , g1 as (select * from g)
   --join by lag
   , j as (select g.id, g.wave, g.time_gps as time_gps_0, g1.time_gps as time_gps_1,
st_makeline(g.the_geom_xxxSRIDxxx, g1.the_geom_xxxSRIDxxx) as geom from g left join g1 on
g.wave = g1.wave and g.id = g1.id and g.xid = g1.xid - 1)
   --calculate time lag, length
   , f as (select *, extract(epoch from time_gps_1 - time_gps_0) as duration_s, st_length(geom)
as length_m from j)
   --calculate MPH, exclude links that were more than 10 minutes
   --select *, (length_ft / 5200) / (duration_s / 3600) as mph_gis from f where duration_s < 600
   select * from f;
   --spatial index
   create index idx_gps_segments on xxxSCHEMAxxx.gps_segments using gist(geom);"

   sqlx <- gsub("xxxSCHEMAxxx", schema, gsub("xxxSRIDxxx", mysrid, sql))

   O <- dbGetQuery(conn = conn, statement = sqlx)
}



# an RLE based function for getting "trips", creates a vector with trip numbers

# some "diff" values
z <- c(30,29,30,31,200,30,29,33,80,20,500,30,29,35,70,900)
# construct the vector using as.numeric(
f.rle <- function(x, tol=150){
   # recode to ones and zeros based on tolerance. 1 part of a run, 0 if after a big time gap
   x.recode <- ifelse(x<tol,1,0)
   # make a RLE object
   x.rle <- rle(x.recode)
   # was the last value 0? If so that means the last record was obtained after a long gap.
   last.rle.val <- x.rle$values[length(x.rle$values)]
   if(last.rle.val==0){
      x.rle$values[length(x.rle$values)] <- 2
   }
   # which RLE values are zeros?
   which.zeros <- which(x.rle$values==0)
   # and which are ones?
   which.ones <- which(x.rle$values>0)
   # which positions come after the zeros?
   after.zeros <- which.zeros + 1
```

```
   # add 1 to the lengths that came after the zeros (increase the "run" lengths by 1)
   x.rle$lengths[after.zeros] <- x.rle$lengths[after.zeros] + 1
   # pull out only the "run" lenghts
   x.rle$lengths <- x.rle$lengths[which.ones]
   # make a sequence of run values
   x.rle$values <- seq(1:length(x.rle$lengths))
   inverse.rle(x.rle)
}


# "create table gps.grid as (SELECT row_number() over() as gid,
(ST_PixelAsPolygons(ST_AddBand(ST_MakeEmptyRaster(30000,30000,483681,10315,5000),
'8BSI'::text, 1, 0), 1, false)).geom as geom);
#
# CREATE OR REPLACE FUNCTION ST_CreateFishnet(
#       nrow integer, ncol integer,
#       xsize float8, ysize float8,
#       x0 float8 DEFAULT 0, y0 float8 DEFAULT 0,
#       OUT "row" integer, OUT col integer,
#       OUT geom geometry)
#    RETURNS SETOF record AS
# $$
# SELECT i + 1 AS row, j + 1 AS col, ST_Translate(cell, j * $3 + $5, i * $4 + $6) AS geom
# FROM generate_series(0, $1 - 1) AS i,
#     generate_series(0, $2 - 1) AS j,
# (
# SELECT ('POLYGON((0 0, 0 '||$4||', '||$3||' '||$4||', '||$3||' 0,0 0))')::geometry AS cell
# ) AS foo;
# $$ LANGUAGE sql IMMUTABLE STRICT;
#
# SELECT row_number() over() as gid, *
# FROM ST_CreateFishnet(6, 6, 5000, 5000, 434350, 10315) AS cells;
#
# create table gps.grid as SELECT row_number() over() as gid, * FROM ST_CreateFishnet(9,10,
5000, 5000, 434350, 10315) AS cells;
# update gps.grid set geom = st_setsrid(geom, 32636);
# alter table gps.grid alter COLUMN geom set data type geometry(polygon, 32636);
#
# create table gps.lines as select id, st_makeline(the_geom_32636 order by
index)::geometry(linestring, 32636) group by id order by id;
#
#
# SELECT create_parent('public.events', 'created_at', 'time', 'daily');
#
```

```
# "
#
#
#
#
# "
# CREATE OR REPLACE FUNCTION measurement_insert_trigger()
# RETURNS TRIGGER AS $$
# BEGIN
#    IF ( NEW.logdate >= DATE '2006-02-01' AND
#         NEW.logdate < DATE '2006-03-01' ) THEN
#        INSERT INTO measurement_y2006m02 VALUES (NEW.*);
#    ELSIF ( NEW.logdate >= DATE '2006-03-01' AND
#            NEW.logdate < DATE '2006-04-01' ) THEN
#        INSERT INTO measurement_y2006m03 VALUES (NEW.*);
#    ...
#    ELSIF ( NEW.logdate >= DATE '2008-01-01' AND
#            NEW.logdate < DATE '2008-02-01' ) THEN
#        INSERT INTO measurement_y2008m01 VALUES (NEW.*);
#    ELSE
#        RAISE EXCEPTION 'Date out of range.  Fix the measurement_insert_trigger() function!';
#    END IF;
#    RETURN NULL;
# END;
# $$
# LANGUAGE plpgsql;
# ""

# an example of the speed threshold algorithm

library(here)
library(sqldf)
library(RPostgreSQL)
library(sqldf)
library(sf)

projectdir <- here::here()
source(file.path(projectdir, "tools", "setup.R"))

# write one GPS data set to the db
db_process <- function(){
   source('/projects/boda/tools/qstarz_to_postgis_partition.R')
   maketablestructure_gps(conn = dbconn, schema = "algo")
   make_child_tables(schema = "algo")
```

```
   qstarz.to.db.allfiles(conn = dbconn, gpsfiles = gpsfiles, schema = "algo", overwrite = FALSE)
   #qstarz.to.db.afile(fname = "/projects/boda/data/gps/safeboda/0002_CS0001.csv", conn =
boda, schema = "algo")
}

# create a table to hold trip numbers
mtt <- function(){
   O <- dbGetQuery(conn = dbconn, statement = "drop table if exists algo.trips;
      create table if not exists algo.trips (id text, time_gps timestamptz, trip boolean, tripnum
integer);
      create index idx_trips_id on algo.trips using btree (id);
      create index idx_trips_time_gps on algo.trips using btree (time_gps);
      create index idx_trips_tripnum on algo.trips using btree(tripnum);")
}


# id="cs0065"; speed_thresh = 5; fast_rec_count_thresh = 10; slow_rec_count_thresh = 45;
fast_is_slow_rec_count_thresh = 180
f_algo <- function(id, speed_thresh = 5, fast_rec_count_thresh = 10, slow_rec_count_thresh =
45, fast_is_slow_rec_count_thresh = 180, plot_out = FALSE, write.db = FALSE){
   message("#####\nprocessing ", id)
   x <- dbGetQuery(conn = dbconn, statement = sprintf("select index, time_gps, speed_kmh
from algo.gps where id = '%s' order by time_gps;", id))
   #record parameters
   x$params <- sprintf("%s;%s;%s;%s", speed_thresh, fast_rec_count_thresh,
slow_rec_count_thresh, fast_is_slow_rec_count_thresh)

   # encode speed threshold
   x$speed_gt_thresh <- ifelse(x$speed_kmh > speed_thresh, TRUE, FALSE)

   # RLE to get slow/fast chunks
   rle0 <- rle(x$speed_gt_thresh)

   # recode chunks that were fast but with fewer than 10 records as being slow -- may require
some fiddling to get the interval right
   rle0$values <- ifelse(rle0$values & rle0$lengths < fast_rec_count_thresh, FALSE, rle0$values)

   # expand
   x$trip1 <- pass1 <- inverse.rle(rle0)

   # now there should be some larger chunks of slow, so redo the RLE
   rle1 <- rle(pass1)

   # if there are any chunks that are slow and shorter than 45 s recode those as fast
```

```
  rle1$values <- ifelse(!rle1$values & rle1$lengths < slow_rec_count_thresh, TRUE,
rle1$values)

  # expand
  x$trip2 <- pass2 <- inverse.rle(rle1)

  # if there are any chunks that are fast and shorter than 3 minutes, recode them as stops
  rle2 <- rle(pass2)
  rle2$values <- ifelse(rle2$lengths & rle2$lengths < fast_is_slow_rec_count_thresh, FALSE,
rle2$values)

  # expand
  x$trip3 <- pass3 <- inverse.rle(rle2)

  # trip numbers
  rlex <- rle(pass1)
  rlex$values <- ifelse(rlex$values, cumsum(rlex$values), NA)
  x$tripnum1 <- inverse.rle(rlex)

  rlex <- rle(pass2)
  rlex$values <- ifelse(rlex$values, cumsum(rlex$values), NA)
  x$tripnum2 <- inverse.rle(rlex)

  rlex <- rle(pass3)
  rlex$values <- ifelse(rlex$values, cumsum(rlex$values), NA)
  x$tripnum3 <- inverse.rle(rlex)

  if(plot_out){
    plotter(x, id)
  }

  # update the db
  if(write.db){
    message("writing to db")
    x2 <- data.frame(id, time_gps = x$time_gps, trip = x$trip3, tripnum = x$tripnum3)
    O <- dbGetQuery(conn = dbconn, statement = "drop table if exists pg_temp.gpsalgo;")
    O <- dbWriteTable(conn = dbconn, name = c("pg_temp", "gpsalgo"), value = x2, row.names
= FALSE)
    sql <- sprintf("delete from algo.trips where id = '%s';
    insert into algo.trips select * from pg_temp.gpsalgo on conflict do nothing;", id)
    O <- dbGetQuery(conn = dbconn, statement = sql)
  }

  return(x)
```

```
}

f_algo.many <- function(plot = TRUE, write.db = TRUE){
   # get the algo.gps IDs
   gpsids <- dbGetQuery(conn = dbconn, statement = "select distinct id from algo.gps order by
id;")
   for(i in gpsids$id){
      #message(i)
      x <- f_algo(id = i, plot_out = plot, write.db = write.db)
   }
   message("updating GPS data with trip ids")
   O <- dbGetQuery(conn = dbconn, statement = "update algo.gps as g set trip = t.trip from
algo.trips as t where g.id = t.id and g.time_gps = t.time_gps;")
   O <- dbGetQuery(conn = dbconn, statement = "update algo.gps as g set tripnum = t.tripnum
from algo.trips as t where g.id = t.id and g.time_gps = t.time_gps;")
}

simulator <- function(id){
   # simulate
   # speed thresholds
   sxt <- seq(from = 4, to = 10, by = 2)
   #sxt <- c(4, 10)
   # slow thresh
   lxt <- seq(from = 8, to = 20, by = 2)
   #lxt <- c(5, 15)
   # fast thresh
   fxt <- seq(40, 120, by = 20)
   #fxt <- c(40, 120)
   # recode short bursts of fast to slow thresh
   fsxt <- seq(from = 60, to = 300, by = 60)

   nruns <- length(sxt) * length(lxt) * length(fxt) * length(fsxt)

   M <- list()

   counter <- 0
   for(i in sxt){
      for(j in lxt){
         for(k in fxt){
            for(l in fsxt){
               # increment the counter
               counter <- counter + 1
               message(paste(counter, "of", nruns))
               # run the algo
```

```
                X <- f_algo(id = id, speed_thresh = i, fast_rec_count_thresh = j,
slow_rec_count_thresh = k, fast_is_slow_rec_count_thresh = l)
                message(paste(i, j, k, l))
                # row names
                nm <- sprintf("%s;%s;%s;%s", i, j, k, l)
                # add this run to the list
                M[[counter]] <- X
                names(M)[counter] <- nm
            }
        }
    }
  }

  assign("M", M, envir = .GlobalEnv)

  # summary stats of speed
  message("collapsing data")
  lss <- do.call(rbind, lapply(M, function(q) sqldf("select params, count(*) as n_records, trip3,
avg(speed_kmh) as mean_speed, stddev(speed_kmh) as sd_speed from q group by trip3,
params;")))
  row.names(lss) <- NULL

  # write to CSV
  write.csv(x = lss, file = sprintf("www/output/%s_sensitivity_params.csv", id), row.names =
FALSE)
}

# trip lines for recoding
triplines <- function(){
  # see tools/triplines.sql

  # and tools/trippoints.sql
}

# recode short duration trips that are also short length as stops
trippoints <- function(){

  # pull the data for RLE
  dat <- dbGetQuery(conn = dbconn, statement = "select * from algo.trippoints_prelim order
by id, time_gps;")

  # recode trips if it was originally a trip but now seems too short (< 3 minutes and < 300 m)
  dat2 <- sqldf("select *,
                 case when trip and duration_s < 180 and length_m < 300 then false
```

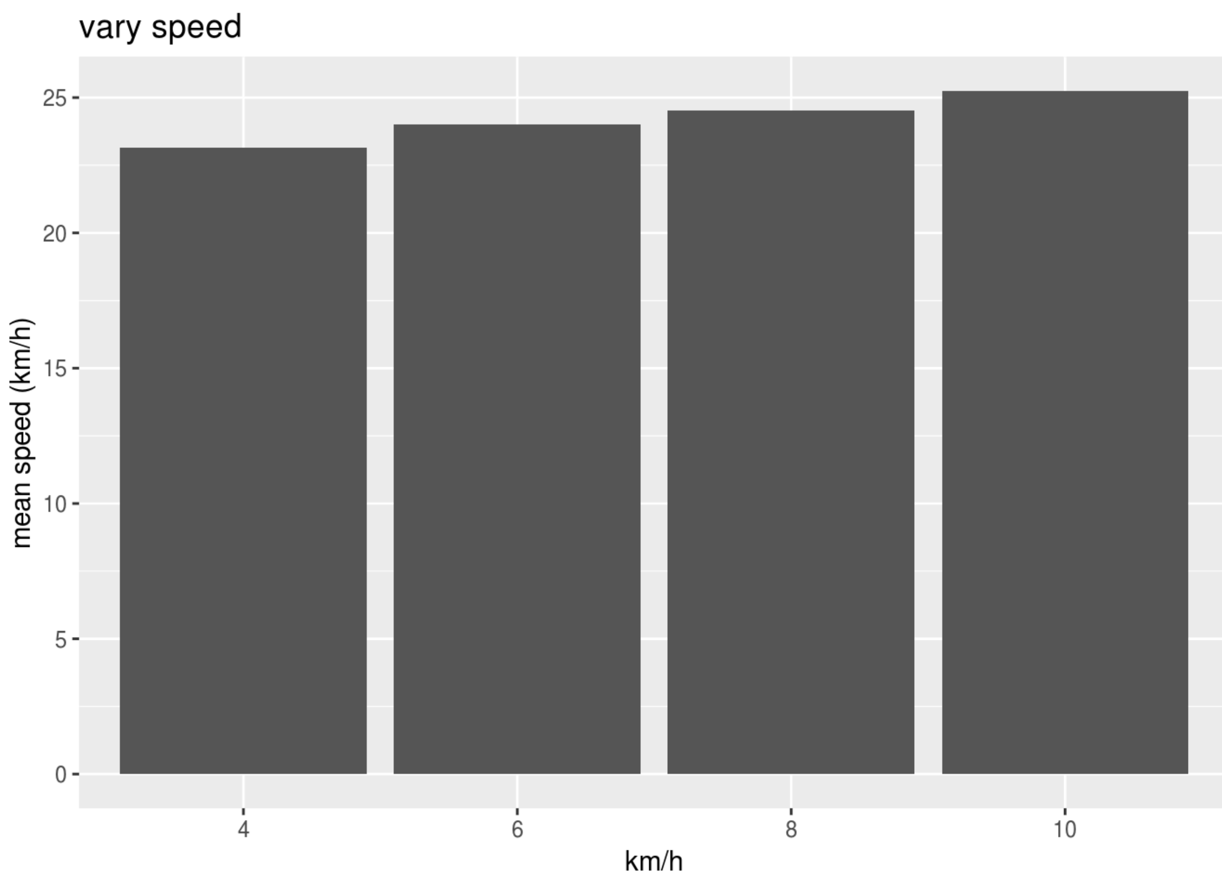
```
            else trip end as newtrip from dat order by id, time_gps")

  # RLE for trip number
  rlex <- rle(dat2$newtrip)
  rlex$values <- 1:length(rlex$lengths)
  # invert the RLE
  dat2$newtripnum <- inverse.rle(rlex)

  O <- dbGetQuery(conn = dbconn, statement = "drop table if exists algo.tripfinal;")
  O <- dbWriteTable(conn = dbconn, name = c("algo", "tripfinal"), value = dat2, row.names =
FALSE)

  O <- dbGetQuery(conn = dbconn, statement = "create index idx_algo_tripfinal_timegps on
algo.tripfinal using btree(time_gps);")

  # make new stop point and trip line data


}

alpha = 0.7
col1 <- adjustcolor(3, alpha.f = alpha)
col2 <- adjustcolor(4, alpha.f = alpha)
col3 <- adjustcolor(5, alpha.f = alpha)
plotter <- function(x, id){
  # plot
  message("plotting ", id)
  pdf(file = sprintf("/projects/boda/www/timeline_graphs/%s.pdf", id), width = 100, height = 8)
  # lines
  plot(x$time_gps, x$speed_kmh, xaxt = "n", ty="l")
  # segments
  segs1 <- sqldf("select tripnum1, min(time_gps), max(time_gps) from x where trip1 group by
tripnum1 order by tripnum1;")
  segs2 <- sqldf("select tripnum2, min(time_gps), max(time_gps) from x where trip2 group by
tripnum2 order by tripnum2;")
  segs3 <- sqldf("select tripnum3, min(time_gps), max(time_gps) from x where trip3 group by
tripnum3 order by tripnum3;")
  # plot segments
  segments(x0 = segs1$min, x1 = segs1$max, y0 = 2, col = col1, lwd=4, lend=1)
  segments(x0 = segs2$min, x1 = segs2$max, y0 = 4, col = col2, lwd=4, lend=1)
  segments(x0 = segs3$min, x1 = segs3$max, y0 = 6, col = col3, lwd=4, lend=1)
  # axis
  at <- unique(round_date(x$time_gps, "10 mins"))
  axis.POSIXct(side = 1, x = x$time_gps, at = at, las = 2)
```

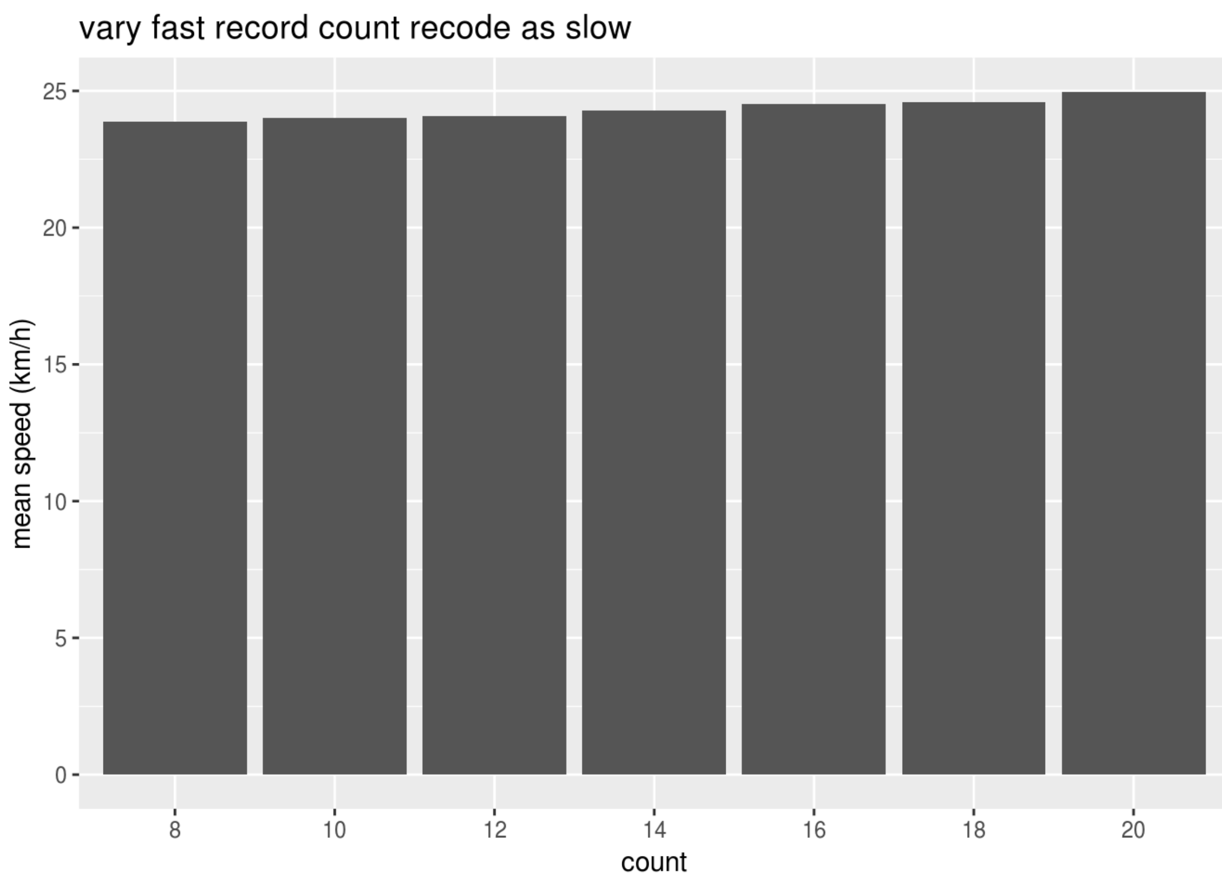
```
   # legend
   legend(x = "topleft", legend = c("raw data", "pass1", "pass2", "pass3"), lwd = c(1, 4, 4, 4),
col=c(1, 3, 4, 5))
   dev.off()
}

# distance to stage
library(here)
library(sqldf)
library(RPostgreSQL)
library(RSQLite)
library(sf)

projectdir <- here::here()
source(file.path(projectdir, "tools", "setup.R"))

# add the distance column
O <- dbGetQuery(conn = dbconn, statement = "alter table algo.gps add column
dist_to_stage_m float;")

# distance query
O <- dbGetQuery(conn = dbconn, statement = "update algo.gps as g set dist_to_stage_m =
st_distancespheroid(s.geom, g.the_geom_4326, 'SPHEROID[\"WGS
84\",6378137,298.257223563]') from
            (select * from gis.subjects join gis.stages using (stage_id)) as s;")


drop table if exists algo.segments;
create table algo.segments as
with
--GPS data
g as (select distinct on (id, time_gps) id, time_gps, tripnum, the_geom_32636, row_number()
over (order by id, time_gps) as index from algo.gps where trip)
--render table
, a as (select id, index, the_geom_32636 as g0, tripnum, time_gps as time_gps_0 from g)
--render table again for self-join
, b as (select id, index, the_geom_32636 as g1, tripnum, time_gps as time_gps_1 from g)
--join
, j1 as (select a.id, a.tripnum, time_gps_0, time_gps_1, time_gps_1 - time_gps_0 as duration,
   st_makeline(g0, g1)::geometry(linestringm, 32636) as geom
   from a
   join b
   on (a.id = b.id and a.index + 1 = b.index and a.tripnum = b.tripnum))
--calculate speed for segment
```

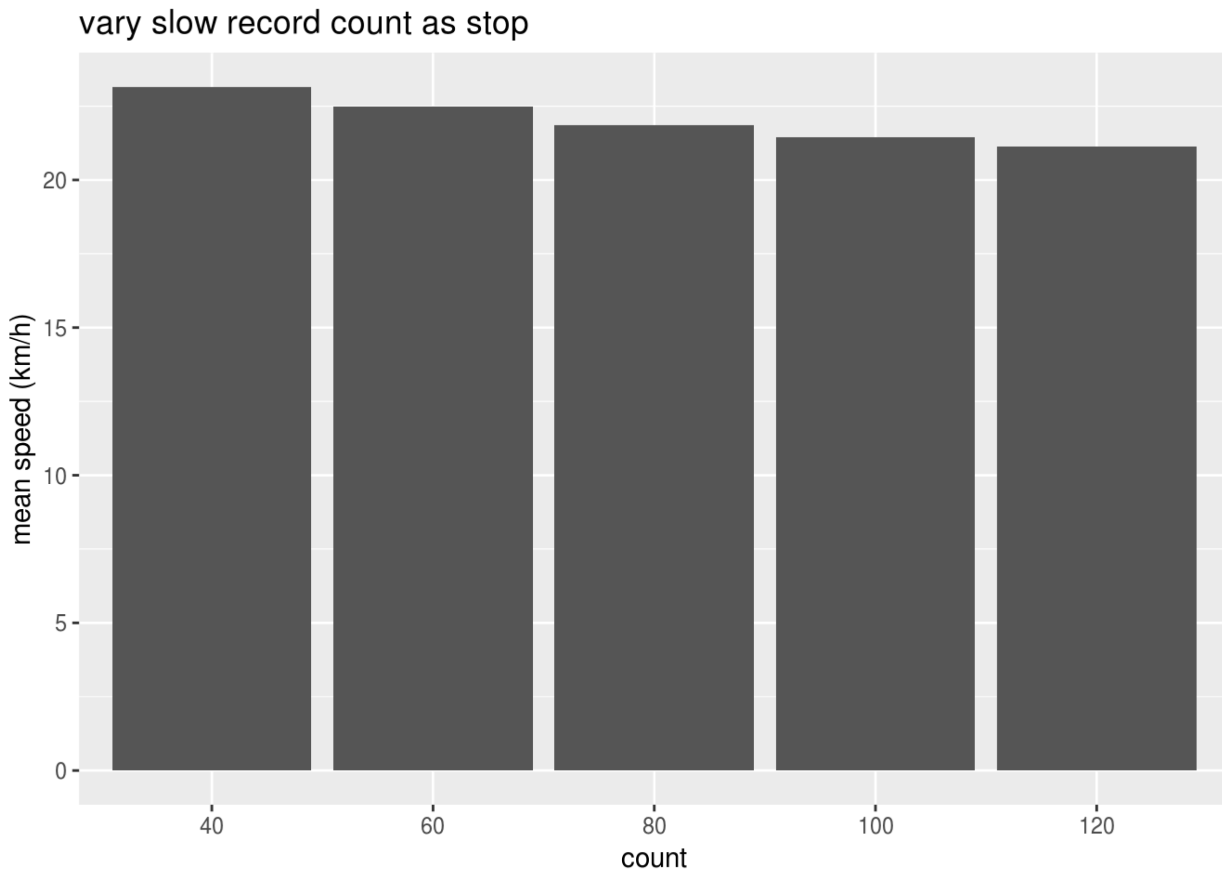
```
, j2 as (select id, tripnum, time_gps_0, time_gps_1, extract(epoch from duration) as duration_s,
geom, st_length(geom) / extract(epoch from duration) * 3.6 as kph from j1)
--drop segments with long duration (these are from where stops were removed)
select * from j2 where duration_s < 45;

drop table if exists algo.tripsum;
create table algo.tripsum as
--point level data
with p as (
select id, tripnum
  , st_length(st_makeline(the_geom_32636 order by time_gps)) as triplength_m
  , min(time_gps) as time_start
  , max(time_gps) as time_end
  , extract(epoch from max(time_gps) - min(time_gps)) as duration_s

  , min(dist_to_stage_m) as min_dist_to_stage_m
  , max(dist_to_stage_m) as max_dist_to_stage_m
  , avg(dist_to_stage_m) as mean_dist_to_stage_m
  , stddev(dist_to_stage_m) as sd_dist_to_stage_m

  from algo.gps
  where trip --and id = 'cs0276'
  group by id, tripnum
  order by id, tripnum)
--segment level data
, s as (select id, tripnum
  , min(kph) as min_speed_kph
  , max(kph) as max_speed_kph
  , avg(kph) as mean_speed_kph
  , stddev(kph) as sd_speed_kph

  from algo.segments
  group by id, tripnum
  order by id, tripnum)

--segment time over 50 kph
, s50 as (select id, tripnum,
  sum(duration_s) as duration_s_above_50_kph
  from algo.segments
  where kph > 50

  group by id, tripnum
  order by id, tripnum)
```

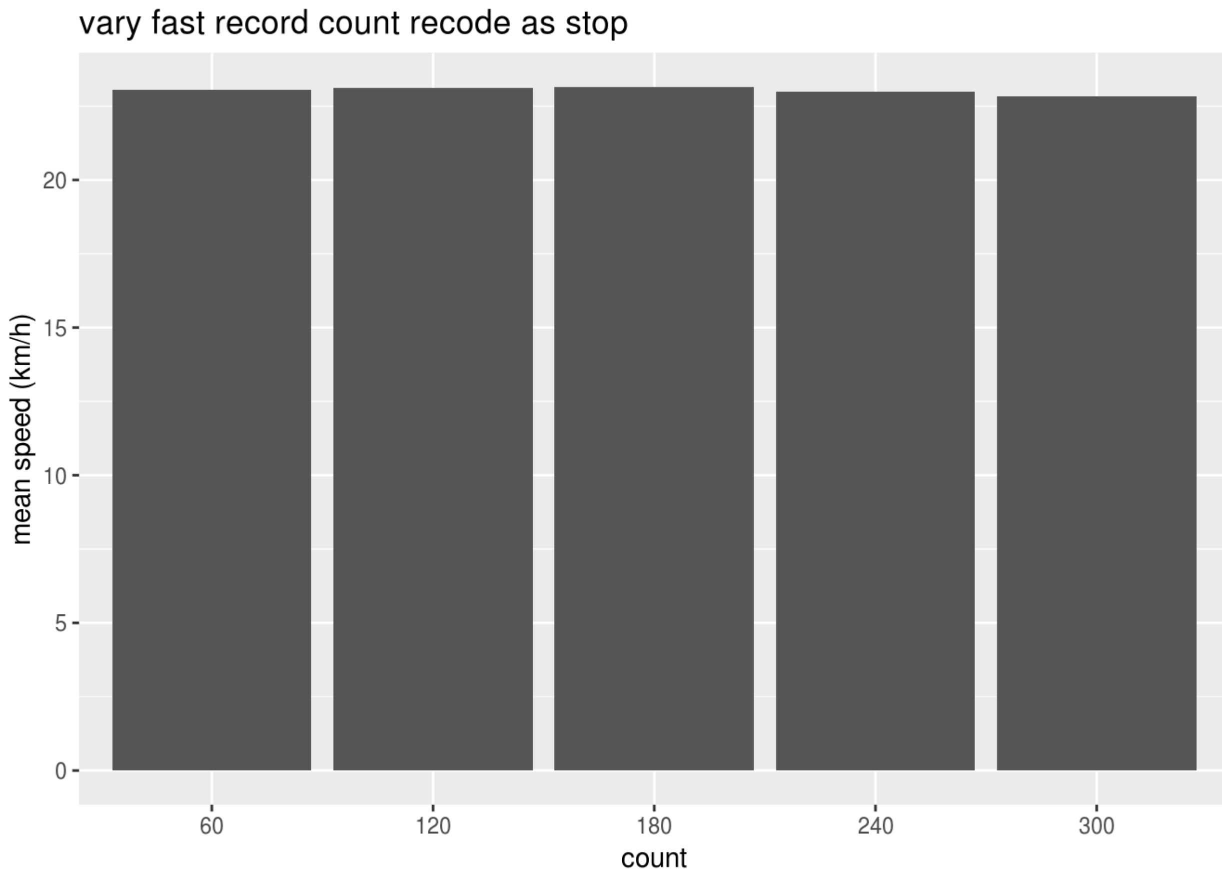
```
--join
, j as (select * from p full join s using(id, tripnum) full join s50 using (id, tripnum) order by id,
tripnum)

select * from j;
```

Figures S3-S6: The following graphs show the effect on mean trip speed of varying one algorithm parameter while holding the others constant at the selected value

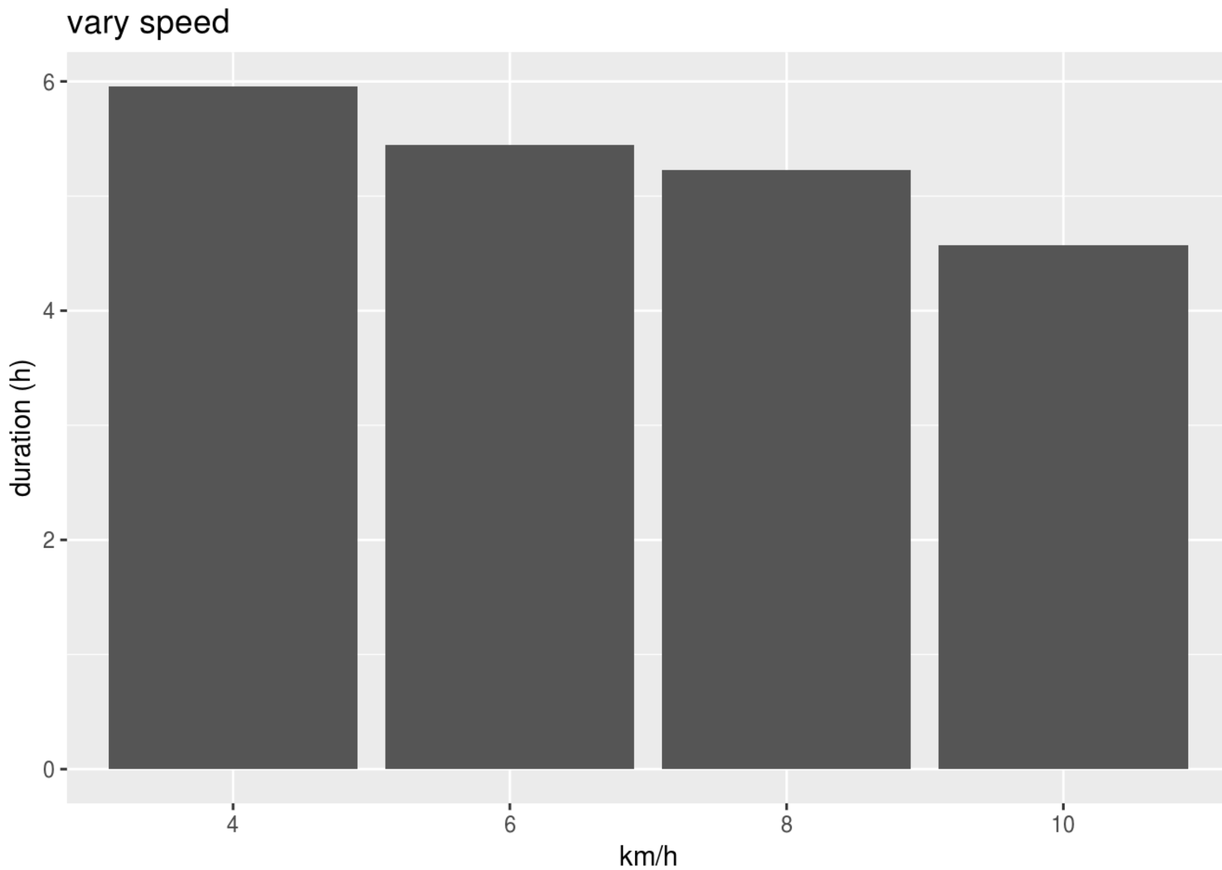


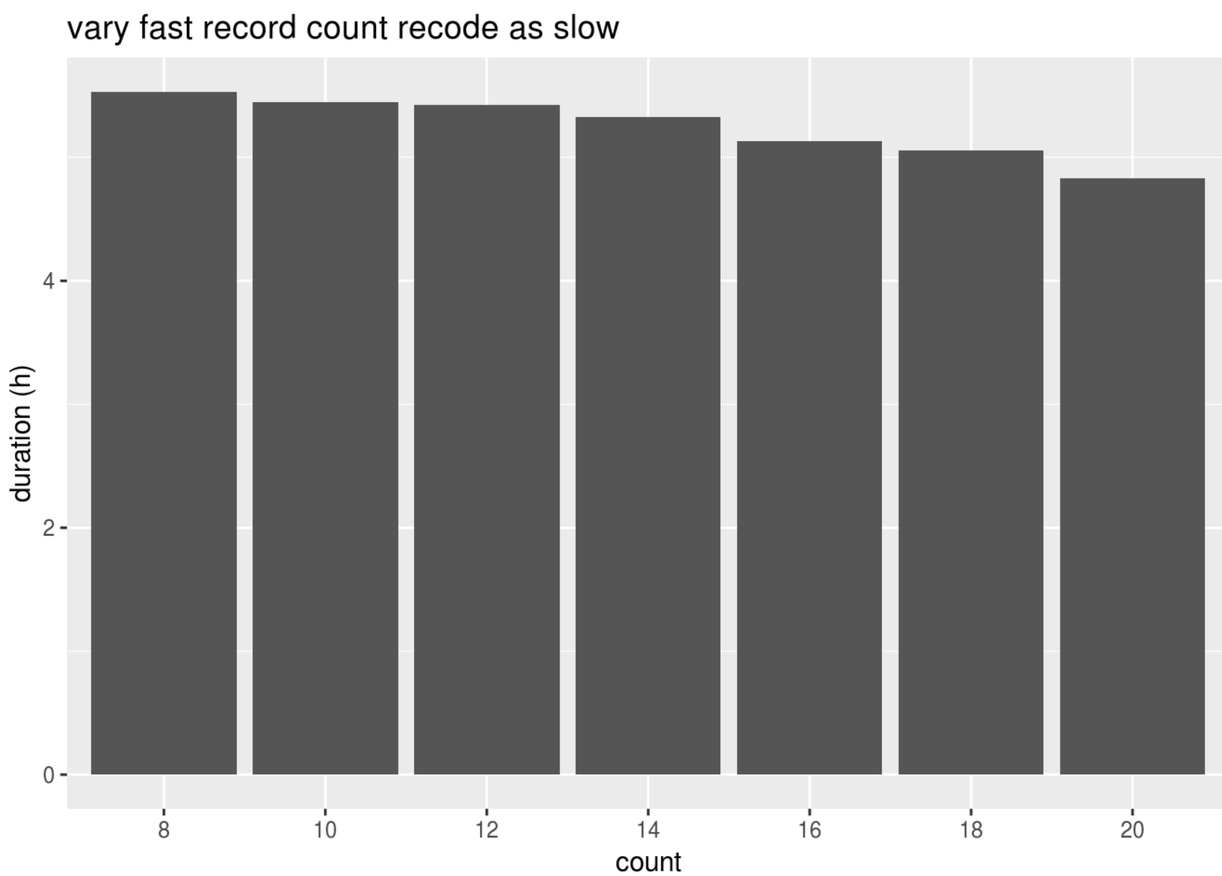

vary fast record count recode as slow
mean speed (km/h)
0
5
10
15
20
25
8
10
12
14
16
18
20
count

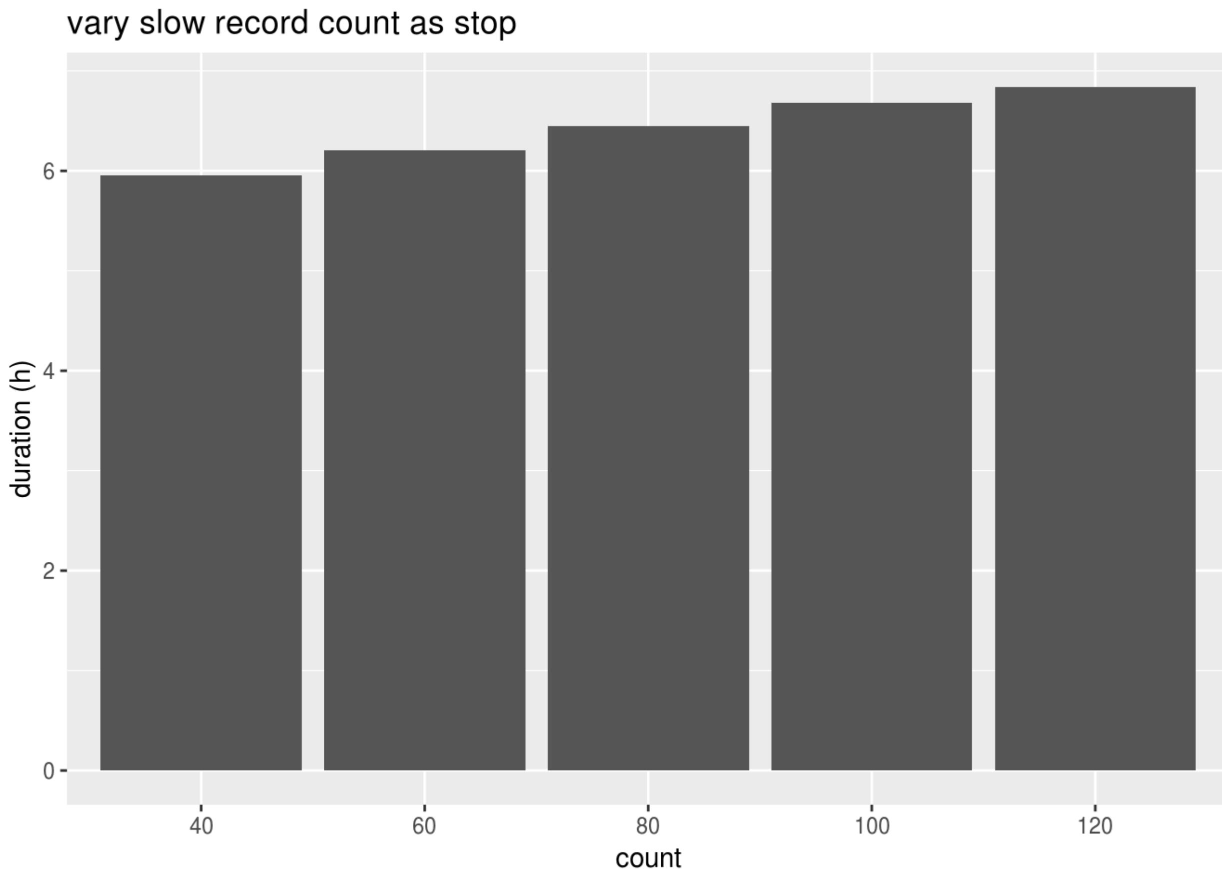

vary slow record count as stop
mean speed (km/h)
0
5
10
15
20
40
60
80
100
120
count

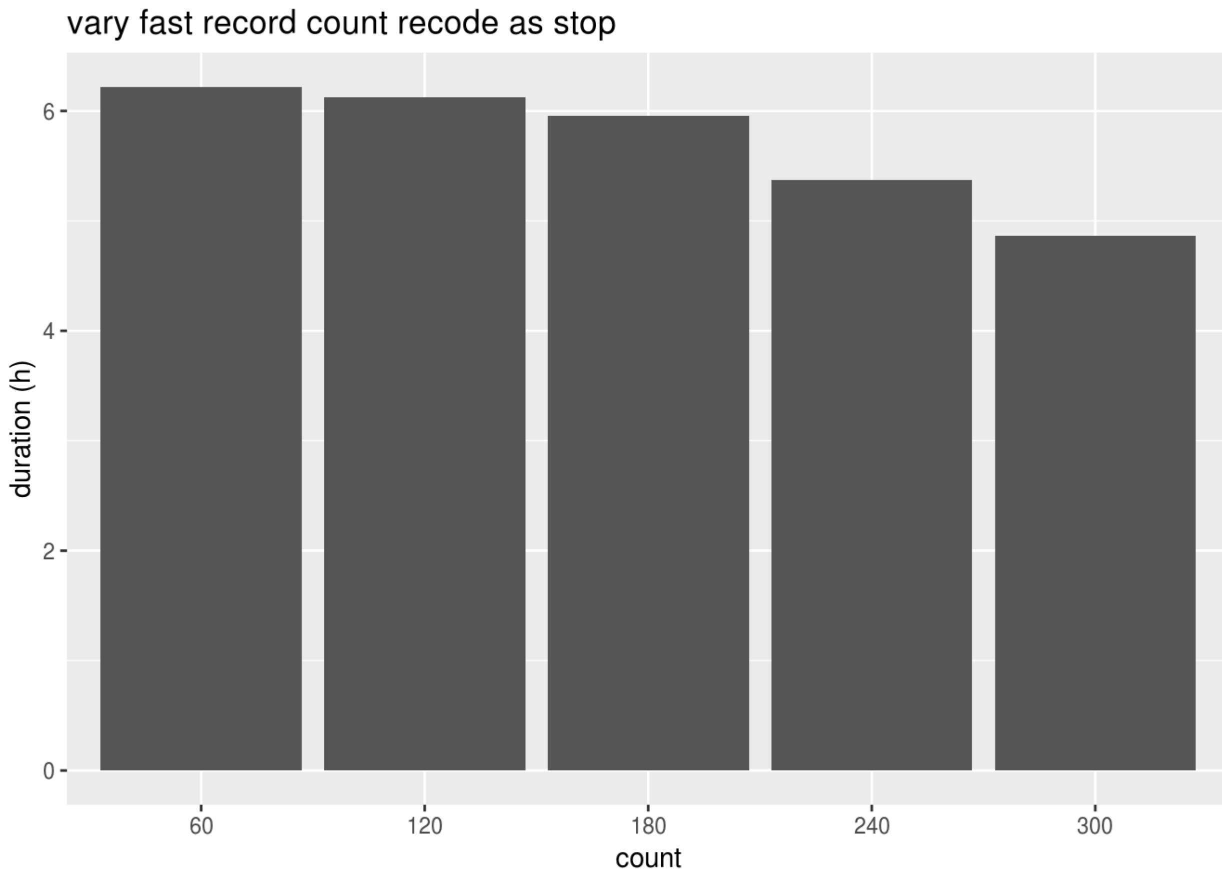

vary fast record count recode as stop
mean speed (km/h)
20
15
10
5
0
60
120
180
240
300
count

Figures S7-S10: The following graphs show the effect on total trips hours of varying one algorithm parameter while holding the others constant




vary fast record count recode as slow
duration (h)
0
2
4
8
10
12
14
16
18
20
count


vary slow record count as stop
duration (h)
0
2
4
6
40
60
80
100
120
count



vary fast record count recode as stop
duration (h)
0
2
4
6
60
120
180
240
300
count