Journal of Public Economics 95 (2011) 1438–1446

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

Heckle and Chide: Results of a randomized road safety intervention in Kenya☆

James Habyarimana [a,*], William Jack [b]

[a] Georgetown University, Public Policy Institute, Washington, DC, USA
[b] Georgetown University, Department of Economics, Washington, DC, USA

ARTICLE INFO

Article history:
Received 23 August 2010
Received in revised form 15 June 2011
Accepted 24 June 2011
Available online 27 July 2011

Keywords:
Road safety

ABSTRACT

We report the results of a randomized field experiment aimed at improving the safety of long-distance minibusses or *matatus* in Kenya. Our intervention combines evocative messages aimed at motivating passengers to speak up against bad driving with a lottery that rewards *matatu* drivers for keeping the stickers in place. Independent insurance claims data were collected for more than 2000 long-distance *matatus* before and after the intervention. Our results indicate that insurance claims fell by a half to two-thirds, from a baseline annual rate of about 10%, and that claims involving injury or death fell by 60%. While we are unable to identify the mechanism(s) underlying this effect, the intervention is more cost effective in reducing mortality than other documented public health interventions.

1. Introduction

This paper reports the results of a field experiment aimed at improving road safety in the developing world. The specific context is that of long-distance road transportation services in Kenya, where it is popularly believed that otherwise rational young males are transformed, Jekyll-and-Hyde-like, into irrational death-seekers when they occupy the driver's seat of a minibus, or *matatu*. Our intervention combines the placement inside vehicles of stickers with messages aimed at motivating passengers to speak up against bad driving with a lottery that rewards *matatu* drivers for keeping the stickers in place. The intervention appears to be extremely cost-effective in terms of reducing mortality, out-performing virtually all other documented public health interventions.

Long distance transportation services in much of the developing world account for a significant share of road traffic injuries and fatalities, which in turn constitute a large and increasing share of both deaths and the disease burden in the developing world. The World Health Organization (2004) reported that 1.2 million people died from road traffic injuries in 2002, 90% in low- and middle-income countries, about the same number as die of malaria. In addition, between 20 and 50 million people are estimated to be injured or disabled each year. Road traffic accidents constitute the largest share, 23%, of deaths due to injury, nearly twice as many as the 14% due to war and violence combined. And they are projected to grow in importance (Lopez et al., 2006), and will be twice as deadly as malaria by 2030 (Mathers and Loncar, 2006).[1,2]

Many interventions to reduce road accidents have been undertaken in developed economies, including programs to reduce the volume of driving, to improve the safety features of road networks, and to enforce traffic regulations more effectively.[3] Publicity campaigns have focused on educating road users, and some have employed shock therapy to get their message across. For example, an advertising campaign in New

☆ We gratefully acknowledge the financial support of the Center for Global Development and the Safaricom Foundation, and thank Channa Commanday and Bright Oywaya of ASIRT-Kenya, the Kenyan branch of the Association for Safe International Road Travel, an international NGO. We thank Mr. Tom Gichuhi of the Association of Kenyan Insurers, senior executive officers of four large Kenyan insurance companies, and executive officers of the 21 matatu savings and credit cooperatives who assisted us in this project. We also thank the Editor and two anonymous referees, as well as Nada Eissa, David Evans, Luca Flabbi, Garance Genicot, Vijaya Ramachandran, Roger Lagunoff and Tavneet Suri for discussions, and seminar participants at Georgetown, the World Bank, and the Kenya Medical Research Institute. We also acknowledge the pro-bono contributions of George Wanjohi and Saracen Media in Nairobi, and John Wali and volunteers from Junior Achievement Kenya. We thank Lauren Marra for excellent research assistance. Finally we thank Philomena Wanjiru, David Gitahi, Asman Wesonga Suleiman and Nadeem Karmali for their tireless and professional work in leading our team of 20 field workers in implementing the study. All errors are our own.

* Corresponding author.
E-mail addresses: jph35@georgetown.edu (J. Habyarimana), wgj@georgetown.edu (W. Jack).

0047-2727/$ – see front matter
doi:10.1016/j.jpubeco.2011.06.008

[1] Country level data are generally less reliable. Odero et al. (2003) suggest that fatality rates in Kenya are extremely high with 7 deaths from 35 road crashes every day, and that the impact of prevailing interventions is dismal. According to a Ministry of Health Report, in 1996 traffic accidents were the third leading cause of death after malaria and HIV/AIDS (Government of Kenya, 1996). More recent estimates suggest that over 3000 individuals died in road traffic related incidents in 2008 (Ministry of Transport, Government of Kenya, 2008).

[2] Road accidents affect the elite as well as the poor. Recent example include the death in March 2009 of the wife of Zimbabwe's prime minister, Morgan Tsvangirai, the serious injury of then future Kenyan president Mwai Kibaki during the election campaign of 2002, and the involvement of former Kenyan president Daniel arap Moi in a serious road accident in 2006.

[3] A comprehensive review of such interventions can be found in World Health Organization (2004), Chapter 4.

Zealand aimed at reducing speeding and drunk-driving, and encouraging the use of safety belts, was found to have an impact on road deaths (Guria and Leung, 2004). Fewer studies of interventions in developing countries exist and while the estimated results of these studies are not causal, measured effects are large. The introduction of speed bumps at certain accident hot-spots in Ghana was associated with a 35% reduction in accidents and a 55% reduction in fatalities (Afukaar et al., 2003). Bishai et al. (2008) found that higher intensity police patrols were associated with a 17% reduction in accident rates in Uganda. Perhaps more creatively, in Bogotá, Colombia, mimes were used to ridicule pedestrians and drivers who flaunted traffic rules.[4]

The intervention we study was simple and cheap: stickers with evocative messages aimed at passengers and encouraging them to "stand up, speak up" were randomly assigned to just over half of 2276 recruited vehicles. High rates of compliance were ensured by running a monthly lottery among drivers of participating treatment *matatus*, who could win up to 5000 Kenyan Shillings (about $60, or roughly one week's wages) if their vehicle was found to have all stickers intact upon inspection by our field staff. Our main outcome data were collected independently from four insurance companies that together cover more than 90% of sample vehicles, and who were unaware of our intervention at the time it took place. We use insurance claims data for treatment and control vehicles in the two year window bracketing the insertion of the stickers. We identify an impact on driver behavior that is both statistically significant and economically large. Our intent-to-treat estimates indicate that the intervention is associated with a reduction in insurance claims rates of about a half, from a projected counterfactual annual claims rate of about 10%. Our instrumental variables estimate of the average treatment effect on the treated yields an even higher estimate of the impact among compliers. Furthermore, we find that this result is largely due to a reduction in claim events where the driver was at fault. We also document a large reduction in claims involving injury or death. Whether these results are due to the stickers, the lottery (which, as was explained to the drivers, was not meant to reward safe driving explicitly), or some combination of the two, cannot be discerned from the available data.

The rest of the paper is organized as follows. Section 2 describes the intervention, experimental design, data and empirical strategy. We present the results in Section 3, and conclude in Section 4.

2. Intervention

In a pre-intervention survey of passengers we found that one third of respondents reported having felt that their life was in danger on a recent *matatu* trip, but that half had never experienced a life-threatening event. Given this heterogeneity in passenger experiences our intervention included a total of five stickers, with both fear stimuli (graphic images of injuries) and simple text messages, as shown in Fig. 1. The stickers, about 11 by 3 in. in size, were placed on the metal panel between a passenger window and the ceiling of the vehicle, ensuring that at least one sticker was within the eye view of each passenger sitting in the main cabin. None were placed in direct view of the driver or the passengers in the front cabin.

2.1. Experimental design

Recruitment was at the individual driver level, with the written support of the cooperatives or "SACCOs"[5] that organize the bus schedules. The major towns among which our sampled *matatus* operated are illustrated in Fig. 2. In all, 21 SACCOs agreed to participate, and just three refused.

[4] This intervention, supported by the Mayor of Bogotá, Antanas Mockus, was not rigorously evaluated, but reportedly enjoyed high levels of popularity (Caballero, 2004).

[5] Savings and credit cooperatives.

At the initial recruitment, participating SACCOs provided us with lists of license plates of vehicles in their fleets from which a random sample of vehicles to be treated was drawn. But in light of our pilot experience, which revealed that vehicle lists were of variable quality, and during which non-participation rates were observed to be reasonably low, we simplified the recruitment protocol and adopted a field-based sampling procedure. Under this protocol, the pilot treatment assignment was maintained but newly recruited *matatus* were assigned to the treatment and control groups based on the final numeric digit of their license plates (odd = treatment, even = control).

In addition, a follow-up survey undertaken soon after the pilot recruitment period found very low rates of sticker retention among treatment vehicles. To address this problem, we implemented a weekly lottery that was to run throughout the study period. Only fully compliant vehicles, those accepting all five stickers at recruitment, were eligible for the lottery. Three randomly chosen winners were inspected by our field staff every week.[6] If an inspected *matatu* was found to have retained all five stickers, the driver would receive a monetary prize: first prize was 5000 KSh (about $US60), second prize was 3000 KSh ($US35), and third prize was 2000 KSh ($US25).[7]

Panels A and B in Table 1 report descriptive statistics of vehicles and drivers respectively, for the treatment and control groups by random assignment. A comparison of first moments suggests that the randomization performed well, there being only one attribute (out of 30) exhibiting a statistically significant difference (at the 5% level) between the two groups. This difference in the share of drivers who reported having had an accident in the last 12 months (third last row of Panel A) — could be of some concern. However, when we examine the same indicator using independent insurance claims data (second last row, Panel A), this difference disappears, suggesting that responses to this question may have been affected by treatment status.[8]

Compliance to the randomized assignment was high but not perfect. Table 2 reports that about 84% of vehicles assigned to the control group complied, and that the same share of those assigned to the treatment group took at least one sticker, with 68.5% taking all five, and 8.0% taking just three (typically the three text-only stickers).

Imperfect compliance to the randomized assignments, either due to driver self-selection or fieldworker error, yielded some statistically significant differences in characteristics by actual treatment status, as reflected in Panel A of Table 3. However, the difference in self-reported accident rates that was significant for random assignment was narrower for actual treatment status. It is conceivable, although not necessarily obvious, that this small narrowing of the difference in self-reported accidents reflects selective adoption of the treatment (even by some assigned to the control) by relatively safe drivers, which would bias the OLS results in favor of finding an effect. To avoid any such bias, our estimates of the average treatment effect rely on intent-to-treat and instrumental variable estimation strategies.

2.2. Data and empirical strategy

In addition to baseline data collected at recruitment, we were granted access to a comprehensive database of claims data from four

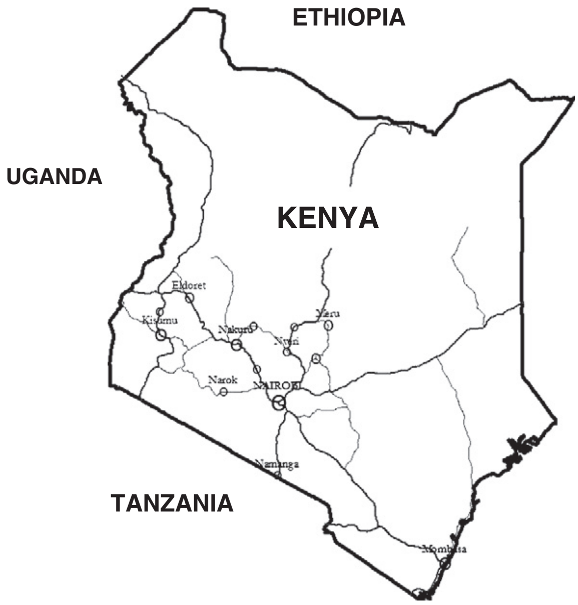
[6] At recruitment, we requested drivers provide us with their cell phone numbers, or a number at which they could be reached. To increase the perceived expected winnings, the treatment group was divided into 5 groups of roughly 200 *matatus* each. Each group's lottery was run every 5 weeks.

[7] Implementing the lottery was challenging, particularly given security concerns in and around the bus stations. The winning license plate numbers were randomly drawn off-site, after which one of our field staff would contact the driver and inspect the vehicle. If it was found to be in compliance, another field staff member would be informed by phone, and would send money via M-PESA, a cell-phone based money transfer system, to the driver. The driver would confirm on the spot receipt of the prize.

[8] Indeed, drivers were administered the recruitment questionnaire *after* they were assigned to the treatment or control group, and those in the treatment group may have been induced to think more about their accident experiences, or even to exaggerate them. In any case, we do not use driver reports as our main outcome variable.

Fig. 1. Stickers inserted in treated *matatus.*

insurance companies that cover over 90% of long-distance *matatus* in our sample (see Panel B in Table 3). There are three possible concerns associated with the use of insurance claims data as an outcome measure. Firstly, as claims are filed by drivers, owners, or passengers there is likely measurement error in observed accidents due to selective filing of claims. However, we do not believe that the decision to file a claim is systematically correlated with randomized assignment to treatment since insurance companies were unaware of which vehicles were participating in the study and the owners and drivers were unaware of the source of our outcome data. While the resulting classical measurement error has implications for precision, it should not bias our results. The second issue is that we do not have access to insurance claims data on the entire sample of vehicles in the study. Panel B of Table 3 compares selected vehicle, trip and driver characteristics by whether we have claims data or not. While there are no statistically significant differences in vehicle and trip



Note: The figure above shows the location of major towns that comprise the main destinations served by matatus sampled for this intervention. The main road network linking these destinations is also shown.

Fig. 2. Major towns served by sampled long-distance *matatus*.

characteristics across these two groups, we observe two significant differences in the driver characteristics: drivers of *matatus* for which we have claims data are more likely to have secondary schooling and are more likely to operate more than one *matatu*. It is difficult to say whether these differences compromise the representativeness of the claims sample. Finally, we do not observe whether or how soon a vehicle associated with a claim resumes service after the claim-generating event. Our simplifying assumption that each *matatu* continues to operate after an accident biases the result against us finding an effect of the intervention.

The claims data were collected for the period January 2006 through May 2009. We use annualized insurance claims rates as an outcome measure, as well as evidence based on our own coding of the description of the accidents such as whether the driver was at fault, and whether injuries or fatalities occurred.

We are interested in estimating the average causal effect of the sticker intervention on the outcomes outlined above. Using outcome information before and after sticker insertion we estimate the following specification:

$$Y_{it} = \alpha + \beta_1 P_{it} + \beta_2 TR_i + \beta_3 P_{it} * TR_i + \beta_4 X_{it} + \eta_i + \varepsilon_{it} \quad (1)$$

where Y_{it} represents the annualized claim rate for *matatu* i during period t, P_{it} is an indicator that takes on the value of 1 for all time periods after recruitment and 0 otherwise, and TR_i is an indicator equal to 1 if the *matatu* was 'treated' and 0 otherwise. Finally X_{it} represents a set of covariates that might include the vehicle condition, and driver and route characteristics, and η_i represents unobserved fixed characteristics of the driver, route and vehicle.

The main parameter of interest is β_3 which captures the *net* change in the outcome variable Y_{it} for treated vehicles compared with those in the control group. A negative and significant coefficient indicates a statistically significant decline in the claims rates among treatment *matatus*. This estimate, and the alternatives described below, likely represent lower bounds on the true value of the parameter due to potential spillovers across treatment and control *matatus*. If the effect of the stickers on individual passengers and/or drivers is durable, those who have been exposed to the treatment may affect outcomes on future trips in control *matatus*.

The identifying assumption in order to recover the average treatment effect β_3 requires that unobserved factors captured by $\eta_i + \varepsilon_{it}$ are uncorrelated with treatment status. We implement two identification strategies and a number of robustness checks to establish the validity of our estimates.

First, since Panels A and B of Table 1 confirm the plausibility of the identifying assumption when TR_i corresponds to the random assignment rule, we estimate the intent-to-treat parameter $\hat{\beta}_3^{itt}$. Second, as Table 2 demonstrates, compliance to random treatment assignment is not perfect: 16% of *matatus* in the control arm did not comply with their assignment. In addition, just over 68% of *matatus* assigned to the treatment arm accepted all five stickers, and 16% of them accepted none. Since imperfect and possibly selective compliance can dilute the estimated effects of assignment to treatment, we use an instrumental variables strategy to estimate the average treatment effect on the treated, using the indicator for random assignment to the treatment group as an instrument for receiving the stickers. The resulting estimator, $\hat{\beta}_3^{att}$, represents the local average treatment effect of the stickers for the group of vehicles whose treatment status is affected by random assignment. Both estimators assume parallel trends in claims rates across both the treatment and control groups. While random assignment should assure this, we present a number of robustness checks below that allow for differential trends.

3. Results

3.1. Effects on insurance claims

A visual summary of the results is presented in Fig. 3, in which the trajectories of claims events per 1000 *matatus* are shown, separately for vehicles assigned to treatment and control, from the first quarter of 2006 to the second quarter of 2009. The horizontal axis in Fig. 3 measures calendar time. Given the considerable lag of 3–6 months in the digital recording of claims, our data for the first two quarters of 2009 are incomplete.[9] The vertical line indicates when recruitment of *matatus*

[9] Our outcome data is defined using the date of the accident and not the date that the claim was filed, processed or digitally recorded.

Table 1
Vehicle and driver characteristics by random assignment.

	Control	Treatment	Difference P-value
Panel A: Vehicle characteristics			
Odometer reading	356,506	361,386	0.612
	(7236)	(6350)	[0.288]
	[327,266]	[343,603]	
Seating capacity	14.52	14.52	0.995
	(0.05)	(0.05)	
Proportion use tout	0.45	0.48	0.087
	(0.02)	(0.01)	
Number of weekly trips	20.19	19.60	0.211
	(0.36)	(0.30)	
Average daily distance, kilometers	420.48	414.10	0.433
	(6.14)	(5.33)	
	[400]	[400]	
Proportion with an installed speed governor	1.00	1.00	0.373
	(0.001)	(0.001)	
Share owned by large cooperative (>300 vehicles)	0.49	0.51	0.419
	(0.02)	(0.01)	
Involved in accident in last 12 months, self reported	0.004	0.015	0.008
	(0.002)	(0.004)	
Insurance claim filed in last 12 months before recruitment, (from administrative data)	0.061	0.071	0.355
	(0.008)	(0.007)	
F-stat and p-value of joint test of significance of all vehicle characteristics	1.02		0.415
Number of observations	1006	1155	
Panel B: Driver characteristics			
Has access to phone[a]	0.96	0.98	0.052
	(0.01)	(0.00)	
Owns a phone[a]	0.89	0.91	0.135
	(0.01)	(0.01)	
% less than 30 years old	18.5	16.2	0.612
	(3.4)	(3.0)	
% 30–40 years old	54.8	56.1	0.831
	(4.3)	(4.1)	
% Primary schooling	22.8	26.2	0.494
	(3.5)	(3.5)	
% Secondary schooling	13.9	14.7	0.842
	(2.8)	(2.8)	
% Married	74.8	77.0	0.665
	(3.7)	(3.5)	
Number of children	2.0	2.0	0.918
	(0.1)	(0.1)	
Proportion drivers assigned to one car only	0.72	0.70	0.649
	(0.04)	(0.04)	
Proportion drivers started after recruitment	0.37	0.41	0.515
	(0.04)	(0.04)	
Median driver tenure, days	296	305.5	0.89
F-stat and p-value of joint test of significance of all driver characteristics	0.39		0.95
Number of observations	139	145	

Notes: Standard errors are in parentheses; medians are in square brackets. The table presents mean/median of vehicle characteristics by treatment assignment. The sample is restricted to *matatus* for which information on random assignment is available. 115 *matatus* that could not be matched to the initial assignment list are dropped. The statistics reported in Panel B of the table are based on a random sample of 284 *matatu* drivers who were surveyed 6 months after recruitment.
[a] Statistics reported in these rows are based on the sample of all recruited matatus.

started.[10] Not surprisingly, quarterly claims rates are very noisy, so that before recruitment we observe moderate albeit insignificant differences across the treatment and control groups. While the sign of the differences in claims rates between treatment and control vehicles oscillates before recruitment, it is consistently negative in the post recruitment period. In particular, claims rates for *matatus* assigned to receive the stickers are considerably lower in the quarters after recruitment.

Consistent estimation of β_3 requires that the trend in accident rates in the absence of treatment be parallel across both groups. As the fitted lines in the figure show, a parallel trends assumption is consistent with the pattern of observed outcomes across the two groups.[11] The claims

[10] Recruitment started in February 2008 and was completed in the second quarter of 2008.

[11] In estimating the trend for the control *matatus*, we use information from 2006 to 2008 to determine the trend and avoid the bias inherent in using the incomplete information from 2009. The trend for the treatment group uses all the quarters in 2006 and 2007.

Table 2
Compliance to the intervention.

Number of stickers actually inserted	True assignment (%)	
	Treatment	Control
0	16.1	84.4
1	3.6	0.3
2	3.1	0.2
3	8.0	0.5
4	0.7	0.1
5	68.5	14.5
Total	100.0	100.0

Notes: The table presents the percentage of vehicles assigned to each experimental group with a particular number of stickers. The sample is restricted to *matatus* for which information on random assignment is available. 115 *matatus* that could not be matched to the initial assignment list are dropped.

Table 3
Selected means of vehicles with and without stickers.

Panel A: Selection on observables	No stickers	Stickers	Difference P-value
Has access to phone	0.96	0.99	<0.001
	(0.01)	(0.00)	
Owns a phone	0.87	0.93	<0.001
	(0.01)	(0.01)	
Odometer reading	354,581	363,247	
	(7093)	(6461)	0.367
	[324,568]	[346,064]	[0.173]
Seating capacity	14.56	14.48	0.25
	(0.05)	(0.05)	
Proportion use tout	0.44	0.49	0.044
	(0.02)	(0.01)	
Number of weekly trips	20.00	19.76	0.617
	(0.36)	(0.30)	
Average daily distance, kilometers	418.65	415.63	0.708
	(5.74)	(5.67)	
	[400]	[400]	
Proportion with an installed speed governor	1.00	1.00	0.35
	(0.001)	(0.002)	
Share owned by large cooperative	0.47	0.53	0.006
	(0.02)	(0.01)	
Involved in accident in last 12 months, self reported	0.006	0.013	0.07
	(0.002)	(0.003)	
Insurance claim filed in last 12 months before recruitment	0.062	0.07	0.487
	(0.008)	(0.007)	
Number of observations	1035	1126	
Panel B: Selection on outcome data	**No claims data**	**Claims data**	
Share of vehicles (%)	8.8	91.2	
Vehicle and trip characteristics			
Odometer reading	280,423	322,449	0.362
	(45,015)	(12,620)	
	[199,993]	[292,980]	
Seating capacity	14.56	14.32	0.369
	(0.25)	(0.11)	
Number of weekly trips	19.12	18.68	0.761
	(1.37)	(0.52)	
Average daily distance, kilometers	433.48	422.39	0.783
	(39.25)	(11.14)	
Share owned by large cooperative	0.64		
	0.48		
	0.112		
	(0.1)	(0.03)	
Driver characteristics			
Owns a phone*	0.88	0.88	0.996
	(0.07)	(0.02)	
% Secondary schooling	5.28	15.19	0.026
	(3.95)	(2.13)	
% Married	88.0	74.81	0.063
	(6.63)	(2.71)	
Proportion drivers assigned to one car only	0.84	0.69	0.067
	(0.07)	(0.03)	

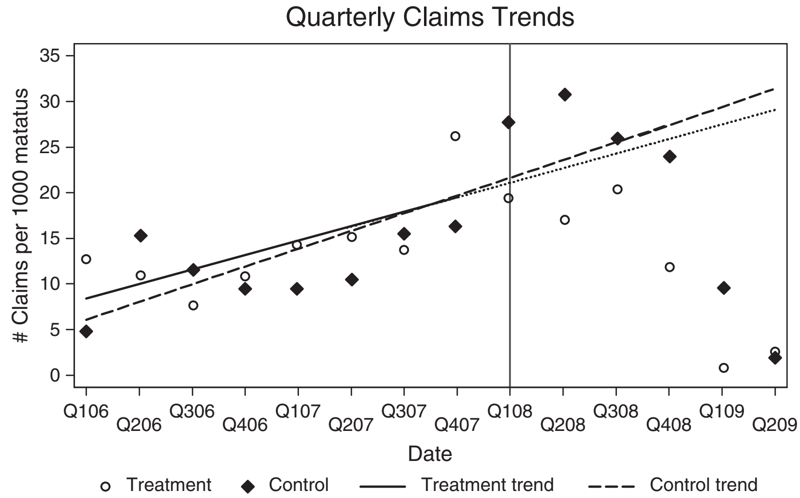
Notes: Standard errors are in parentheses; medians are in square brackets. The P-value in [ ] is drawn from a quantile regression. Panel A presents mean/median of driver/vehicle characteristics by actual treatment status. The sample is restricted to *matatus* for which information on random assignment is available. 115 *matatus* that could not be matched to the initial assignment list are dropped. Panel B presents selected means by whether the insurer of a particular vehicle was one of the four firms that provided claims data.
* Data is drawn from baseline survey.

rate trajectory for the control *matatus* starts a little lower but is marginally steeper so that the trend lines cross. We show below that our results are robust to allowing for group specific trends.

One concern suggested by Fig. 3 is the apparent increase in accident rates for control *matatus* in the first three quarters of 2008. However, consistent with the upward trend in *matatu* claims rates, Fig. 4 shows a similar pattern for larger 30–41 seater vehicles over the same period, suggesting a secular trend in accident rates.

Panel A of Table 4 presents intent-to-treat estimates of the impact of the intervention. The effect on all claims (column (1)) is a large negative and significant reduction in claims rates of five percentage points (p-value<0.01). Including controls for the management group (SACCO) increases the magnitude of this effect only slightly. In 312 of the 362 claims events in our data (about 86%) that we could classify,[12] the *matatu* driver is recorded as being at fault. These claims fell by 4.63 percentage points (p-value<0.01), a 53% reduction, as a result of the intervention (columns (3) and (4)). Similarly, our data include 227 claims with at least one injury or death, and the intervention was associated with a 60% reduction (p-value<0.01) in such accidents from a projected base of 6.65%.

[12] Two claims had no accompanying descriptions that could be used for this coding exercise.



Note: The figure presents the number of insurance claims by quarter between January 1 2006 and May 25 2009. *All* insurance claims are used to construct this figure. Solid and dashed lines represent fitted linear trends for the treatment and control group. We fit a linear trend to all claims for the pre-treatment period for the treatment group (all claims from 2006-2007). The dotted line traces out counterfactual claims for the treatment group. For the control group, we fit a linear trend to all claims from 2006-2008, excluding claims from quarters 1 and 2 of 2009 due to incompleteness. We make the simplifying assumptions that *matatus* continue to operate after a claim event and were operating throughout this period.

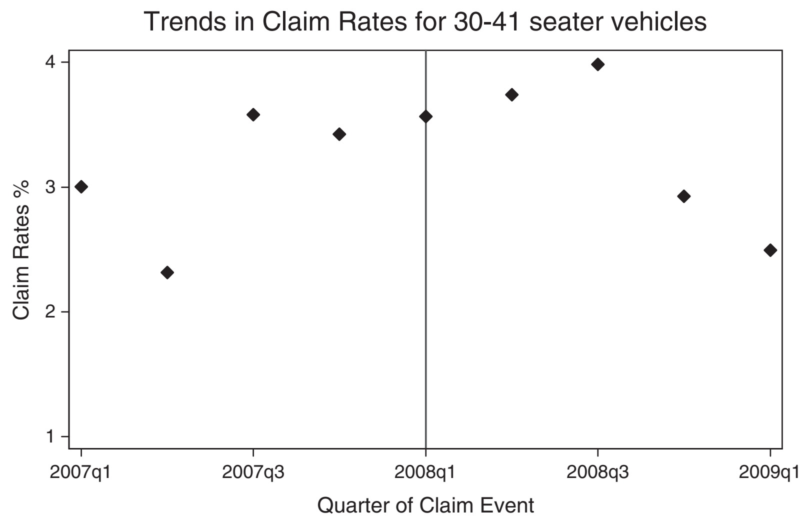
Fig. 3. Quarterly claims trends for recruited sample: January 2006–May 2009.

Panel B presents our instrumental variables results. In columns (1) and (2) we report a 7.3 percentage point reduction in all claims rates amongst those vehicles whose actual treatment status is affected by the instrument. Relative to the projected claims rate, the local average treatment effect suggests a decline in the rate of accidents of as much as 73% associated with the treatment. We observe even larger treatment effects for driver-at-fault and injury–death claims. While the usefulness of some IV results is legitimately questioned in the face of weak instruments and impact heterogeneity, we believe our strong first stage and high compliance rates make this a credible estimate of the impact of the intervention. Among compliers to the instrument, around three-quarters of the accidents that would otherwise have occurred are avoided.

3.2. Robustness checks and cost-effectiveness

In Table 5, we present both linear probability (Panel A) and marginal probit (Panel B) estimates using quarterly data and allowing for seasonality and group specific trends. The results in column (1) in Table 5 are analogous to column (1), Panel A of Table 4 with the sole difference that Table 4 uses annualized claim rates. In columns (2) and (3) we include a secular time trend and calendar quarter fixed effects. Neither of these additions should change the estimated impact of the intervention since the timing of recruitment is orthogonal to the assignment. In column (4) we include a group specific trend which allows the pre-recruitment trends in accident rates to differ across the two groups. The point estimates remain virtually unchanged and indicate large and significant reductions in *quarterly* claims rates of about 1.2 percentage points (Panel A) and about 1 percentage point (Panel B) respectively. In the case of the linear probability model of Panel A, this represents a reduction of about 50% compared with the projected rate.

The cost of the intervention was just under $2 per vehicle for the stickers, and $5 per vehicle per year for the lottery, or a total of $7000 per 1000 vehicles per year. We estimate that the intervention saved about



Note: Figure generated using data on the ratio of the number of 30-41 seater vehicles reporting a claim in a particular quarter to the total number of insured 30-41 seater vehicles in that quarter. Data is drawn from one of the larger insurance companies for which seat class is provided. The vertical line corresponds to the quarter in which the majority of *matatus* were recruited.

Fig. 4. Claims rates for 30–41 seater vehicles that provide long distance transportation services.

Table 4
Regression results.

	All claims		Driver-at-fault claims		Injury/death claims	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Intent-to-treat						
Post	0.029	0.030	0.025	0.026	0.018	0.018
	(0.013)*	(0.012)*	(0.011)*	(0.011)*	(0.009)+	(0.009)*
Assigned to treatment	0.010	0.009	0.011	0.011	0.011	0.011
	(0.010)	(0.011)	(0.010)	(0.010)	(0.008)	(0.008)
Post × Assigned to treatment	−0.050	−0.051	−0.046	−0.047	−0.040	−0.041
	(0.016)**	(0.016)**	(0.014)**	(0.014)**	(0.012)**	(0.012)**
Constant	0.061	0.042	0.052	0.039	0.038	0.036
	(0.008)**	(0.013)**	(0.007)**	(0.012)**	(0.006)**	(0.010)**
Panel B: IV Estimates						
Effect of treatment on the treated	−0.073	−0.075	−0.068	−0.069	−0.059	−0.060
	(0.023)**	(0.023)**	(0.021)**	(0.021)**	(0.017)**	(0.017)**
Controls for SACCO		X		X		X
Observations	4322	4318	4322	4318	4322	4318
R-squared	0.003	0.02	0.002	0.01	0.002	0.01
Mean post recruitment claims rate for vehicles assigned to control group	0.09		0.077		0.055	
First stage: F-stat	2421.33	2364.44				

Notes: Robust standard errors in parentheses. Table reports the estimates of ordinary least squares regression in specifications (1–4) and instrumental variables estimates in specifications (5–6). The dependent variable is the annualized rate of a claim-generating accident for each *matatu* in the sample. We make a simplifying assumption that *matatus* continue to operate after a claim event and were operating throughout the pre- and post-recruitment period. First stage F-stat reports the F-stat of the test of the null that random assignment to treatment does not predict actual treatment status at recruitment. The sample excludes 3% of recruited vehicles for which treatment assignment information could not be reliably established.
* Significant at 5%.
** Significant at 1%.

1200 years of life (see Habyarimana and Jack (2010)). The cost per life year saved is thus about $5.80 (0.8% of per capita GDP) including the lottery costs, and $1.70 (0.2% of per capita GDP) counting only the material costs. Measures of the cost per Disability Adjusted Life Year (DALY) gained would be smaller still once the reduced numbers of injuries were included. The cost-effectiveness of the intervention, including lottery costs, is thus lower than that of childhood vaccination, which at $7 per DALY gained is considered to be among the most cost-effective health interventions available, and it is an order of magnitude lower than virtually all interventions that are considered to be "good buys" for developing countries, such as tuberculosis therapy using the directly observed treatment — short course (DOTS) strategy ($102 per DALY), and improved emergency obstetric care ($127 per DALY).[13]

4. Conclusions

We have presented evidence that a very cheap intervention can alter the behavior of drivers in the context of long distance minibus transportation services in Kenya. Our intent-to-treat estimates suggest that the intervention reduced the number of incidents leading to an insurance claim by at least half. Although the estimated effect is very large, we argue that it is nonetheless plausible, as the intervention provides passengers and perhaps drivers with timely and salient reminders of the consequences of inaction.

A number of alternative mechanisms could underlie our results. Consistent with the intervention's rationale, the stickers could have empowered passengers and legitimized complaints about poor driving. Alternatively, drivers might have anticipated such responses and adjusted their behavior accordingly, even in the absence of equilibrium heckling. On the other hand, drivers may have been directly affected by the stickers, which could have made them aware of the potentially catastrophic outcomes of unsafe driving. And finally, the drivers might have mistakenly understood that only vehicles with claims-free histories would be eligibility for the lottery.

[13] Jamison et al. (2006), and reported in Disease Control Priority Project (2008, Fig. 2).

We lack the observational data on real-time behavior of passengers and drivers to address these issues directly. In an earlier working paper we reported the results of passenger and driver surveys that provided suggestive evidence consistent with passenger empowerment, and we argued that the lottery was unlikely to be generous enough to have

Table 5
Testing robustness of ITT estimates: group specific time trends.

	Dependent variable: Indicator for claim-backed accident in quarter			
	(1)	(2)	(3)	(4)
Panel A: Linear probability model				
Post	0.011	0.006	0.006	0.006
	(0.003)**	(0.004)	(0.005)	(0.006)
Assigned to treatment	0.002	0.002	0.002	0.001
	(0.002)	(0.002)	(0.002)	(0.004)
Post × Assigned to treatment	−0.012	−0.012	−0.012	−0.013
	(0.004)**	(0.004)**	(0.004)**	(0.006)*
Time trend		0.001	0.001	0.001
		(0.000)*	(0.000)	(0.001)
Assigned to treatment × Time trend				0.000
				(0.001)
Constant	0.013	0.009	0.008	0.008
	(0.001)**	(0.002)**	(0.003)*	(0.003)*
Panel B: Marginal probit estimates				
Post × Assigned to treatment	−0.009	−0.009	−0.009	−0.008
	(0.002)**	(0.002)**	(0.002)**	(0.003)*
Calendar quarter fixed effects			X	X
Observations	28783	28783	28783	28783
R-squared	0.00	0.00	0.00	0.00

Notes: Standard errors clustered at the vehicle level are in parentheses. Panel A presents the results of a linear probability estimation of the likelihood of having an accident by quarter. Panel B presents marginal probit estimates for each specification shown above. This estimation uses all insurance claims matched to the experimental sample from January 2006 to May 2009. The sample is restricted to *matatus* for which information on random assignment is available. 115 *matatus* that could not be matched to the initial assignment list are dropped. We make a simplifying assumption that *matatus* continue to operate after a claim event and were operating throughout this period.
* Significant at 5%.
** Significant at 1%.

induced the very large changes in driving behavior implied by the observed reduction in claims. Nonetheless, additional work will be needed to definitively shed light on the underlying mechanisms, and the implications for the design of similar interventions in other environments.

References

Afukaar, F.K., Antwi, P., Ofosu-Amah, S., 2003. Pattern of road traffic injuries in Ghana: implications for control. Injury Control and Safety Promotion 10, 69–76.

Bishai, D., Asiimwe, B., Abbas, S., Hyder, A., Bazeyo, W., 2008. Cost effectiveness of traffic enforcement: case study from Uganda injury prevention. Injury Prevention 14, 223–227.

Caballero, Maria Cristina, 2004. Academic Turns City into a Social Experiment: Mayor Mockus of Bogota and his Spectacularly Applied Theory. Harvard University Gazette. (March 11).

Disease Control Priorities Project, 2008. Using cost-effectiveness analysis for setting health prioritiesMarch http://www.dcp2.org/file/150/DCPP-CostEffectiveness.pdf.

Government of Kenya, 1996. Ministry of Health. Health Information System, 1996 Report. Government Printers, Nairobi.

Government of Kenya, 2008. Ministry of Transport. Statistical Abstract 2008. Government Printers, Nairobi.

Guria, Jagadish, Leung, Joanne, 2004. An evaluation of a supplementary road safety package. Accident Analysis and Prevention 36 (5), 893–904.

Jamison, D.T., Breman, J.G., Measham, A.R., Alleyne, G., Claeson, M., Evans, D.B., Jha, P., Mills, A., Musgrove, P., 2006. Disease Control Priorities in Developing Countries, 2nd Edition. Oxford University Press, New York.

Habyarimana, James, Jack, William, 2010. Heckle and Chide: Results from a Randomized Road Safety Intervention in Kenya. Georgetown University, Draft.

Lopez, A., Mathers, C., Ezzati, M., Jamison, D., Murray, C., 2006. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. Lancet 367, 1747–1757.

Mathers, C., Loncar, D., 2006. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Medicine 3 (11), 2011–2030.

Odero, W., Khayesi, M., Heda, P.M., 2003. Road traffic injuries in Kenya: magnitude, causes and status of intervention. Injury Control and Safety Promotion 10 (1–2), 53–61.

World Health Organization, 2004. World Report on Road Traffic Injury Prevention. WHO, Geneva.