Yellow taxis have fewer accidents than blue taxis because yellow is more visible than blue

Teck-Hua Ho[a,b,1], Juin Kuan Chong[c], and Xiaoyu Xia[d]

[a]Office of the Deputy President (Research & Technology), National University of Singapore, Singapore 119077; [b]Haas School of Business, University of California, Berkeley, CA 94720; [c]National University of Singapore Business School, National University of Singapore, Singapore 119245; and [d]Department of Decision Sciences and Managerial Economics, Chinese University of Hong Kong Business School, Chinese University of Hong Kong, Shatin, NT, Hong Kong

Is there a link between the color of a taxi and how many accidents it has? An analysis of 36 mo of detailed taxi, driver, and accident data (comprising millions of data points) from the largest taxi company in Singapore suggests that there is an explicit link. Yellow taxis had 6.1 fewer accidents per 1,000 taxis per month than blue taxis, a 9% reduction in accident probability. We rule out driver difference as an explanatory variable and empirically show that because yellow taxis are more noticeable than blue taxis—especially when in front of another vehicle, and in street lighting—other drivers can better avoid hitting them, directly reducing the accident rate. This finding can play a significant role when choosing colors for public transportation and may save lives as well as millions of dollars.

car color | road safety | data science | transportation science | sensory perception

Accidents involving public transport are common and cause significant economic losses as well as loss of human life. Applying statistical analysis to a unique and comprehensive dataset we establish that a change in color can avert a significant number of taxi accidents, leading to a reduction in economic losses. Specifically, analysis of a complete set of accident records from the largest taxi operator in Singapore, which uses yellow and blue taxis, shows that yellow is safer than blue because yellow is more noticeable, with the result that potential accidents are avoided by other drivers' timely responses.

Yellow has been a popular color for taxis since 1907, when the Chicago Yellow Cab Company chose the color based on a survey conducted at the University of Chicago. The survey showed that yellow was the most noticeable color, which would make it easy for potential passengers to spot a yellow taxi in the sea of mass-produced black cars prevalent at the time (until 1914, "Japan Black" was the only paint color that would dry fast enough to be used in Ford's mass-production process). More than a century later, it turns out that yellow was a wise choice, not only for potential passengers but also for actual passengers because yellow taxis seem to have fewer accidents than blue taxis.

Although there is anecdotal evidence on higher accident rates for dark-colored vehicles, few papers have empirically established a strong causal link between color and accident risk. For example, Lardelli-Claret et al. (1) and Furness et al. (2) found that dark-colored cars have a significantly higher risk than light-colored cars of being involved in crashes with serious injuries.* However, insufficient data prevented both studies from accounting for different base rates in the car population. The comprehensive dataset and quasi-experimental design in our study allow us to conduct a clean test of the color effect.

To test whether there was a causal relationship between the color of a taxi and the number of accidents the taxi had we analyzed two datasets from the largest taxi company in Singapore: (*i*) a 36-mo, aggregate accident-record dataset (January 2012 to December 2014) for all taxis and (*ii*) a dataset on a randomly chosen sample of 20% (3,341 drivers) of the company's drivers, which includes 3 mo of data (April to June 2014) on their daily driving records, and 30 mo of data (January 2012 to June 2014) on their taxi contracts, accident records (if any), and basic

demographic characteristics. These two datasets include millions of observations on the company's drivers and taxis, and accidents involving these taxis. The data from both datasets have been anonymized and are available in Datasets S1–S6.

The company uses yellow or blue for all its regular taxis (approximate colors are shown in Fig. 1).[†] The colors are the remnants of a 2002 merger that took place between two taxi companies, one of which used yellow and the other, blue. The company owns ∼16,700 taxis in a ratio of one yellow to three blue (1y:3b), which translates to 4,175 yellow taxis and 12,525 blue ones. These account for 60% of the ∼27,800 taxis in Singapore.[‡]

To control for the difference in the number of taxis used by the company (1y:3b), we calculated a normalized accident rate using the average number of accidents that occurred per 1,000 taxis

*In ref. 1 the authors assumed that the color frequencies of cars involved in collisions were identical to the base rates of colors in the general car population. If car color does matter, less-accident-prone colors will be involved in fewer collisions and hence will be underrepresented in the data. The paper concluded that white and yellow were associated with a slightly lower risk of being passively involved in collisions, and the protective effect of these colors was greatest under worsening visibility. Furness et al. (2) studied all car drivers on public roads in the Auckland region of New Zealand between April 1998 and June 1999. The authors compared 571 car crashes involving hospital admissions with 588 randomly sampled cars to determine the effect of car color on the risk of being in a serious car crash. They found a significant reduction in risk in silver cars compared with white cars, but a significant increase in risk in brown cars. They found that risk in yellow, grey, red, and blue cars was not significantly different from that in white cars.

[†]The company has two fleets, each with its own color. However, one team manages the recruitment, training, and allocation of drivers to taxis for both fleets, using a common pool of drivers. We noticed that the two fleets also have a few white limousines for hire, but we could not obtain the exact number of limousines for hire in each month. The company informed us that both fleets have a similar number of white limousines. Based on the data we have, 95.1% of car accidents in the dataset occurred in either yellow or blue taxis, and 4.9% of the accidents occurred in white limousines. We exclude accidents involving white limousines in our analysis to minimize measurement error.

[‡]The Land Transport Authority of Singapore manages taxi quotas for all taxi companies and issues monthly reports on the number and distribution of taxis belonging to all taxi companies in the country. Naturally, the total number of taxis fluctuates over time; 27,800 was the average number of registered taxis in the second quarter of 2014, which matches the timing of our driving-record data.

Fig. 1. Approximate colors used for taxis.

per month. Analysis revealed that in 33 of the 36 mo, yellow taxis were involved in fewer accidents than blue taxis (Fig. 2).

On average, yellow taxis were involved in 6.1 fewer accidents per 1,000 taxis per month—65.6 compared with 71.7 for blue taxis, a statistically significant difference with a P value < 0.0001 (Fig. 3 and Table 1).

What could account for this notable difference in the accident rate? Perhaps yellow taxis were driven less frequently? Fare structure and monthly rent for taxis of both colors were the same, so there was no difference in economic incentive to motivate different driving behavior.

Perhaps safer drivers preferred driving yellow taxis, or yellow taxis attracted safer drivers? For example, Newman and Willis (3) found that drivers of certain colors of cars were more likely to receive speeding tickets. To rule out this hypothesis, we requested from the company a supplementary dataset with the average speed of yellow and blue taxis every hour over a 1-wk period, giving 168 average-speed pairs. We found that the difference in average speed between the two colors almost always fell in the range of ± 1 km/h (Fig. S1).

We also learned from the company that all drivers were hired using the same recruitment system and underwent the same training. The drivers were randomly assigned a taxi color, regardless of preference. With this common and random recruitment and assignment system it is fair to assume that reckless drivers would be assigned to taxis of both colors in accordance with the relative frequency of colors (1y:3b).

The Drivers

Perhaps the drivers in the two fleets simply drove differently. We analyzed the demography and driving behavior of a random sample of 3,341 drivers (20% of the total) for 3 mo using 15-s-interval location and status data from their taxis (amounting to more than 150 million data points).[§]

We compared three demographic factors that might be related to driving skill: age, education, and experience (Table S1). There were no discernable differences.

Drivers of yellow and blue taxis drove in a similar fashion; they made the same total number of trips and distributed their working hours (total as well as with passengers on board) in similar patterns (Tables 2 and 3). We also compared the average proportion of time each driver spent each day in the 28 districts and the airport.[¶] The χ^2 tests on the difference in proportions between yellow and blue drivers were not statistically significant (Table S2).

The similar demography and driving behavior between the two groups of drivers further validate randomness in the company's taxi assignment mechanism. We also verified the finding in Fig. 2 with the sample of drivers and found a similar effect due to yellow on the accident rate after adding demographic controls (Table S3).

Results

The Color Conjecture. Because we could not attribute the differing accident rates to differences in driver demography or driving behavior, we looked at physical differences between yellow and blue taxis. Color was the only differentiator because the company
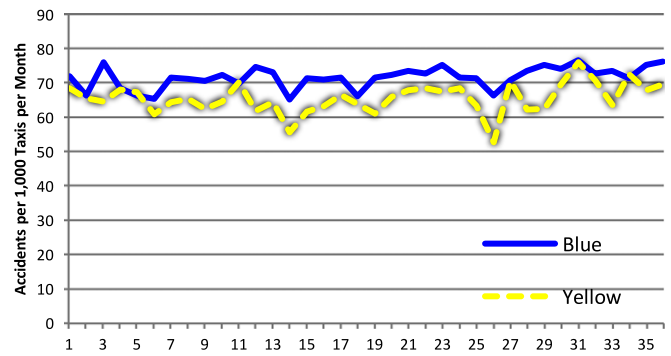

Fig. 2. Thirty-six-month accident trend.

used the same car models and enforced the same maintenance policy for all its taxis.

Of the shades of yellow and blue used by the company yellow is indisputably more visible than blue. Yellow is also a relatively uncommon color for cars, in part because of its long association with taxis.[#] Based on these conditions, we hypothesized that the higher visibility of yellow was directly responsible for the lower accident rate. This higher visibility would make it easier for other drivers to notice a yellow taxi, which would increase the odds that other drivers would have sufficient response time to avoid a potential accident with a yellow taxi.

To test this hypothesis, we looked at the detailed accident reports in the dataset. After each accident occurred the attending police officer noted several facts about the accident, including the nature of the accident and the lighting condition in which it occurred. We tested our color conjecture using these accident reports.

The first test compares the proportion of yellow taxis involved in different types of accidents, focusing on cases where a taxi was unambiguously in or out of the other driver's view. If our conjecture were correct, a yellow taxi would be less likely than a blue taxi to be involved in an accident when the taxi was clearly in the other driver's view.

The accident data obtained from the company included 26 accident descriptors. Of the 26, only two offered unambiguous positions for the two vehicles involved in the accident: *i*) taxi in front, where the taxi was in front of the other vehicle and was hit in the rear by the other vehicle; and *ii*) taxi behind, where the other vehicle was in front of the taxi and was hit in the rear by the taxi.

Table 1. Effect of yellow on the accident rate using
$$A_{im} = \alpha + \beta \, Yellow_{im} + \varepsilon_{im}$$

Variables	Parameter estimates	Robust standard error
Yellow	−0.0061***	(0.0007)
Constant	0.0717***	(0.0004)
Observations	592,296	
R^2	0.0001	

The dependent variable $A_{im} \in \{0,1\}$ is whether driver i had an accident in month m, the key independent variable $Yellow_{im}$ is whether driver i's taxi was yellow in month m, α is the constant, and ε_{im} represents the random error term. The number of observations is based on all registered taxis in the company (ranging from 15,588 to 16,933 per month) in the 36-mo period. R^2 is the coefficient of determination. ***$P < 0.0001$.

[§]The company only keeps Global Positioning System (GPS) records for the most recent 3 mo due to the massive storage requirements for such data. We made the data request in July 2014, so the only GPS records the company could provide were those from April to June 2014.

[¶]Singapore is small and has a total land area of approximately 720 km² (278 square miles). The Singapore government divides the island state into 28 administrative districts. We marked the airport as a separate area because many taxis queue there for passengers.

[#]According to the Land Transport Authority of Singapore, yellow passenger cars account for only 0.98% of the passenger car population (which includes private cars, company cars, driving-school cars, rental cars, and cars that are only used during off-peak hours). Yellow motor vehicles account for just 2.23% of all vehicles (which include passenger cars, taxis, motorcycles and scooters, goods transportation vehicles such as vans and trucks, buses, and tax-exempted vehicles).
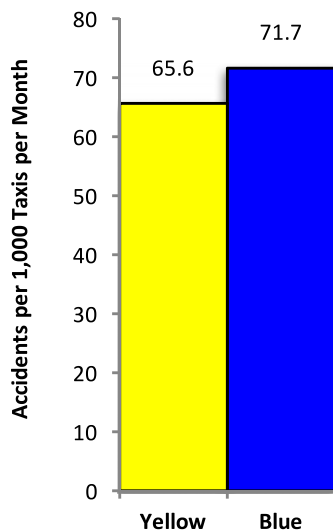
Fig. 3. Difference in accident rate, by color.

Fig. 4. Accident rate by taxi location.

There were 13,925 accidents that matched these two descriptors, which accounted for 33.5% of the total number of accidents in the 36-mo dataset. Because we could clearly identify the positions of the vehicles, we compared the monthly accident rates of yellow and blue taxis attached to these two descriptors. Because the other driver could clearly see the color of the taxi in front, the difference between yellow and blue would be greater in accidents classified as "Taxi in front" (3.4 accidents per 1,000 taxis per month, from 13.4 to 10.0) than in accidents classified as "Taxi behind" (1.6 accidents per 1,000 taxis per month, from 11.4 to 9.8). The results in Fig. 4 confirm our hypothesis; the corresponding regression result is shown in Table S4.

The information on lighting condition provided in the accident reports also allowed us to test the color conjecture. We hypothesized that as long as the higher visibility of yellow over blue was not neutralized (e.g., in total darkness) yellow would remain more noticeable than blue. In fact, yellow's better visibility would be even more advantageous in street lighting because yellow would have a stronger contrast than blue against a dark background (this includes dawn and dusk). We compared accident rates based on the three lighting conditions used in the reports: "street lighting," "daylight," and "no light."‖ As hypothesized, the relative difference in the accident rate was greater in street lighting (4.5 accidents per 1,000 taxis per month, from 27.8 to 23.3) than in daylight (2.0 accidents per 1,000 taxis per month, from 43.7 to 41.7) (Fig. 5). The corresponding regression result is shown in Table S5.

To test the significance in the difference by taxi location and lighting condition, we conducted a difference-in-difference regression analysis. Specifically, we defined the dependent variables to be the difference in the number of accidents between two conditions. The second column in Table 4 shows

the difference between Taxi in front and Taxi behind for the same taxi, and the fourth column shows the difference in the number of accidents between street lighting and daylight for the same taxi. The conclusions shown in Figs. 4 and 5 are supported by these difference-in-difference analyses.

Drivers of Both Colors. A total of 868 drivers (out of the sample of 3,341) drove both yellow and blue taxis. The accident records of this set of "switching" drivers provided a compelling test for the color conjecture. We added driver fixed effects to control for any differences in accident propensity caused by driver characteristics. We performed the analysis based on daily driving records. If accident occurrences are independently and identically distributed across days, we can scale the daily estimate of 0.204 fewer accidents per 1,000 yellow taxis in Table 5 by the average number of days per month (365/12) to get a monthly accident reduction estimate. The resulting monthly reduction is 6.2 accidents per 1,000 taxis when a driver drove a yellow taxi compared with a blue taxi. This is remarkably similar to the figure of 6.1 accidents per 1,000 taxis per month that was obtained earlier using the complete accident dataset.

Discussion

How many accidents can be avoided by simply switching the color of all taxis to yellow? Fig. 3 shows that yellow taxis have 6.1 fewer accidents per 1,000 taxis per month. Table 5 examines data on driver switching behavior and controls for driver fixed effects and shows that yellow taxis had 6.2 fewer accidents per 1,000 taxis per month. We therefore use the more conservative number of 6.1 to evaluate the economic implications of switching the color of all taxis to yellow.

If the company changed the color of its entire fleet of 12,525 blue taxis to yellow, 76.4 fewer accidents would occur per month, or 917 fewer accidents per year. Assuming an average repair cost of SGD 1,000 per car and a downtime of 6 d, we are looking at an annual savings of SGD 2 million.**

Table 2. Driver working behavior

Daily averages	Yellow	Blue
Passenger trips	13.4	13.6
"On-call" trips	2.6	2.8
Work time, min	609	599
Income (SGD)	206	207

Table 3. Breakdown of driver working time

Status	Yellow, min	Blue, min	Yellow, %	Blue, %
On break	141	135	23.2	22.5
Looking for passenger	205	199	33.7	33.2
Passenger on board	196	198	32.2	33.1
Other	67	67	11.0	11.2

‖It is difficult to find places that are not properly lit in the city-state of Singapore. In fact, there were so few accidents recorded in unlit areas that we could not use a statistically significant test.
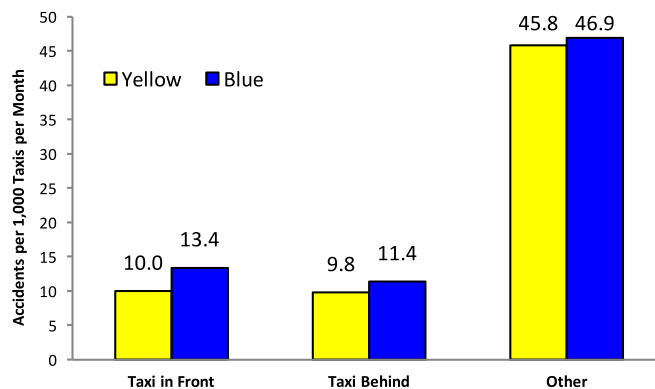
**The taxi company provided an estimated average downtime of 6 d. The annual savings are computed as [(1,000 × 1) + (6 × 200)] × 917, or [(the cost per car × minimum number of vehicles involved in an accident) + (vehicle downtime days × daily earnings)] × number of accidents.
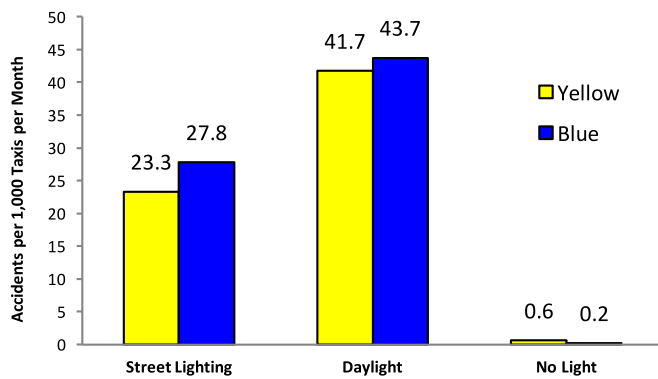
Fig. 5. Accident rate by lighting condition.

Let us evaluate the physical risk to a taxi passenger. Assume a passenger commutes via taxi 5 d/wk between two places that are 10 km apart. The monthly distance traveled is 450 km (10 miles × 2 trips a day × 5 d × 4.5 wk), which is comparable to the average driving distance of 461 km/d that was given by the company. If we assume monthly accident rates of 65.6 (yellow) and 71.7 (blue) per 1,000 vehicles per month, over the course of 40 y a passenger will experience 1.1 accidents in a blue taxi but only one accident in a yellow one, which is a 9% reduction that would please any passenger.

As a thought experiment, we extrapolated economic and accident outcomes for two major cities using the estimate of 6.1 fewer accidents per 1,000 vehicles per month. We start with the approximate number of taxis found in cities that predominantly use colors other than yellow (Table 6). Interestingly, New York City predominantly uses yellow taxis.

These results would be especially noteworthy to smaller taxi companies and to drivers who use their private vehicles as taxis to work for Uber or Lyft. In fact, the growing market of privately owned vehicles serving as taxis may be an ideal test case for the color change because these vehicles are no different in appearance from other privately owned cars, which are not as noticeable as yellow taxis.

Our calculations of economic impact and physical risk produce only the lower bound for the color advantage because we did not calculate the economic impact and physical risk to the other party involved in the accident.

It is an empirical question as to whether the noticeability of yellow would be reduced if the color were to become more common. Fortunately, in this case, even if the company changed the color of its entire fleet of blue taxis to yellow, yellow would remain rare because the company's 16,700 yellow taxis (together with all other yellow vehicles) would only account for about 3.5% of the 1 million vehicles in Singapore, an increase of about 1.2%.

Conclusion

In the early 20th century the color yellow was chosen because of its noticeability, and it is that very noticeability that makes yellow taxis not only easier to see but also safer to ride in than blue taxis. Because yellow is an uncommon color for private cars, it is likely that results similar to those discussed here would also be observed if taxis of other colors were changed to yellow. This will not only reduce the number of accidents that take place but will also reduce the associated loss in economic activity. It could turn out that a simple commercial decision made by the Chicago Yellow Cab company more than a century ago has an inadvertent, positively impactful economic and potentially life-saving outcome that we can adopt and expand on, starting today.

Table 5. Effect of yellow on the accident rate among drivers who drove both yellow and blue taxis using $A_{id} = \alpha_0 + \beta\ Yellow_{id} + \alpha_1 X_{id} + D_i + \varepsilon_{id}$

Variables	Parameter estimates	Robust standard error
Yellow	−0.000204*	(0.000063)
Driving experience in 10 y	−0.002194	(0.002217)
Age in 10 y	0.002699	(0.002204)
Relief driver[†]	−0.001133***	(0.000242)
Driver fixed effects	Yes	
Observations	934,875	
R^2	0.0018	

The dependent variable $A_{id} \in \{0,1\}$ is whether driver i had an accident on day d, the independent variables include whether driver i's taxi was yellow on day d ($Yellow_{id}$), the driver's driving experience and age, whether the driver was a relief driver on that day, and the driver's personal fixed effect dummies. The analysis is done at the daily level to account for drivers switching colors on any day of a month. The number of observations is based on 868 drivers who had active contracts with both yellow and blue taxis. In this specification, one observation is one driver's driving record on 1 d with one type of taxi; the per-day coefficients can be scaled by a factor of 365/12 to convert them to a per-month estimated level for easy comparison with values in the other tables. R^2 is the coefficient of determination. ***$P < 0.0001$, *$P < 0.01$.
[†]Regular drivers rent the taxis directly from the taxi company. These drivers are allowed to let "relief" drivers use their taxis as long as the relief drivers are on the approved list of drivers maintained by the company. On average, regular drivers work 11.5 h/d and relief drivers work 8.5 h/d.

Table 4. Difference-in-difference estimate of the effect of yellow on the accident rate by taxi location and lighting condition using $A_{im}(1) - A_{im}(2) = \alpha + \beta\ Yellow_{im} + \varepsilon_{im}$

Variables	Front (1) – Behind (2)		Street lighting (1) – Daylight (2)	
	Parameter estimates	Robust standard error	Parameter estimates	Robust standard error
Yellow	−0.0018***	(0.0004)	−0.0025*	(0.0008)
Constant	0.0020***	(0.0002)	−0.0159***	(0.0004)
Observations	592,296		592,296	
R^2	0.0000		0.0000	

The dependent variable is the number of accidents when the taxi was in front (or in street lighting) minus the number of accidents when the taxi was behind (or in daylight) for one driver in a given month, the key independent variable $Yellow_{im}$ is whether the taxi was yellow, α is the constant and ε_{im} represents the random error term. The number of observations is based on all registered taxis (ranging from 15,588 to 16,933 per month) in the 36-mo period. R^2 is the coefficient of determination. ***$P < 0.0001$, *$P < 0.01$.

Table 6. Extrapolated reduction in accidents and related savings

City	No. of taxis	Fewer accidents per 1,000 per month	Fewer accidents per 1,000 per year	Savings (SGD, million)
Beijing	65,000	397	4,764	10
London	23,300	142	1,706	3.8

1. Lardelli-Claret P, et al. (2002) Does vehicle color influence the risk of being passively involved in a collision? *Epidemiology* 13(6):721–724.
2. Furness S, et al. (2003) Car colour and risk of car crash injury: Population based case control study. *BMJ* 327(7429):1455–1456.
3. Newman MC, Willis FN (1993) Bright cars and speeding tickets. *J Appl Soc Psychol* 23(1): 79–83.

ECONOMIC SCIENCES

Supporting Information
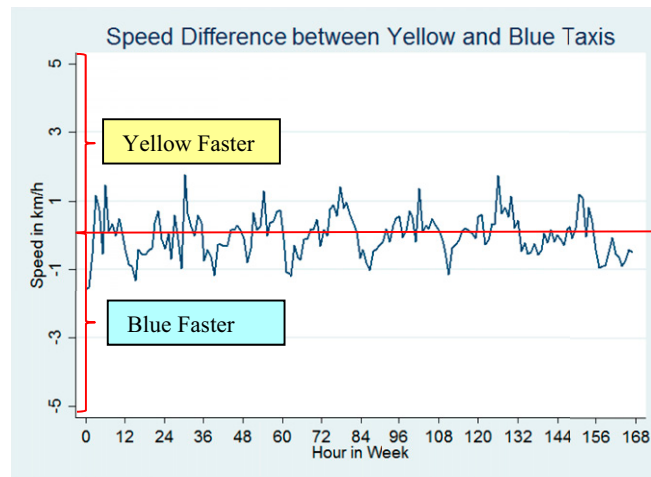
Ho et al. 10.1073/pnas.1612551114

Fig. S1. Driving speed difference (yellow – blue) by hours. The figure is based on the hourly average driving speed of all yellow and blue taxis in a supplementary dataset provided by the company. The *P* value for the paired sample *t* test for mean difference is 0.8498, so we cannot reject the hypothesis that the average driving speed is the same between yellow and blue taxis.

Table S1. Driver demography

Demographic	Yellow, %	Blue, %
Age, y		
<36	4	4
36–40	6	5
41–45	10	9
46–50	16	18
51–55	21	23
56–60	21	23
>60	22	17
Education		
Primary	15	14
Secondary	77	80
Tertiary	9	6
Experience, y		
<3	54	56
3–6	14	13
7–10	12	11
11–15	11	10
16–20	6	6
21–25	2	2
26–30	0	1

The difference in age is insignificant at $\chi^2 = 12.52 < \chi(0.01,6) = 16.81$. The difference in education is insignificant at $\chi^2 = 7.48 < \chi(0.01,2) = 9.21$. The difference in experience is insignificant at $\chi^2 = 4.82 < \chi(0.01,6) = 16.81$.

Table S2. Driving time breakdown by district in percentages for yellow and blue taxis

District	Passenger on board, %		Total time, %	
	Blue	Yellow	Blue	Yellow
1	5.1	5.1	4.7	4.7
2	1.2	1.2	1.1	1.1
3	4.2	4.2	4.8	5.0
4	2.5	2.7	2.5	2.7
5	4.6	4.6	3.9	3.9
6	1.0	1.0	0.9	0.8
7	2.6	2.7	2.6	2.6
8	1.6	1.6	2.1	2.3
9	5.5	5.5	4.4	4.3
10	3.6	3.8	3.0	3.1
11	5.5	5.5	3.3	3.2
12	5.2	5.1	4.7	4.7
13	2.6	2.6	1.8	1.9
14	6.8	6.7	6.4	6.2
15	4.6	4.7	3.4	3.3
16	4.8	4.7	5.4	4.9
17	1.3	1.4	1.7	1.5
18	4.8	4.7	5.6	5.8
19	7.2	7.3	8.0	8.1
20	4.4	4.3	6.0	5.9
21	2.1	2.1	1.4	1.4
22	5.8	6.1	6.2	6.9
23	3.9	3.9	4.8	4.7
24	0.2	0.3	0.2	0.2
25	2.5	2.5	3.3	3.1
26	2.4	2.3	1.2	1.2
27	2.3	2.3	3.5	3.5
28	1.3	1.3	0.9	0.9
29	1.1	1.5	3.0	3.6

The difference in time breakdown by district for passenger on board is insignificant at $\chi^2 = 2.13 < \chi$ (0.01,28) = 48.28. The difference in total time breakdown by district is insignificant at $\chi^2 = 3.54 < \chi$ (0.01,28) = 48.28.

Table S3. Effect of yellow on the accident rate with drivers' demographic characteristics as control variables using $A_{id} = \alpha_0 + \beta\, Yellow_{id} + \alpha_1 X_{id} + \varepsilon_{id}$

Variables	Parameter estimates	Robust standard error
Yellow	−0.000303***	(0.000050)
Driving experience in 10 years	−0.000226***	(0.000046)
Age in 10 years	0.000115***	(0.000035)
Relief driver	−0.001007***	(0.000056)
Years of education	0.000020	(0.000013)
Driver's residence fixed effects	Yes	
Observations	2,862,669	
R^2	0.0002	

The dependent variable $A_{id} \in \{0,1\}$ is whether driver i had an accident on day d, and the independent variables include whether driver i's taxi was yellow on day d, the driver's driving experience and age, whether the driver was a relief driver on that day, and the driver's residence fixed effect dummies (defined by the first two digits of the driver's home zip code). The analysis is done at the daily level to account for drivers switching colors on any day of a month. The number of observations is based on the number of drivers with active contracts that day, among the 3,341 drivers. The per-day coefficients can be scaled by a factor of 365/12 to convert them to a per-month estimated level for easy comparison with the other tables. R^2 is the coefficient of determination. ***$P < 0.0001$.

Table S4. Effect of yellow on the accident rate by taxi location

Variables	Taxi in front		Taxi behind		Other	
	Parameter estimates	Robust standard error	Parameter estimates	Robust standard error	Parameter estimates	Robust standard error
Yellow	−0.0034***	(0.0003)	−0.0016***	(0.0003)	−0.0011	(0.0006)
Constant	0.0134***	(0.0002)	0.0114***	(0.0002)	0.0469***	(0.0003)
Observations	592,296		592,296		592,296	
R^2	0.0002		0.0000		0.0000	

The dependent variable (DV) $A_{im} \in \{0,1\}$ is whether a driver had an accident in a certain location in a given month, and the key independent variable is whether the taxi was yellow. The number of observations is based on all registered taxis (ranging from 15,588 to 16,933 per month) in the 36-month period. R^2 is the coefficient of determination. ***$P < 0.0001$.

Table S5. Effect of yellow on the accident rate by lighting condition

Variables	Street lighting		Daylight		No light	
	Parameter estimates	Robust standard error	Parameter estimates	Robust standard error	Parameter estimates	Robust standard error
Yellow	−0.0045***	(0.0005)	−0.0020**	(0.0006)	0.0004***	(0.0001)
Constant	0.0278***	(0.0002)	0.0437***	(0.0003)	0.0002***	(0.0000)
Observations	592,296		592,296		592,296	
R^2	0.0001		0.0000		0.0001	

The dependent variable $A_{im} \in \{0,1\}$ is whether a driver had an accident under certain lighting conditions in a given month, and the key independent variable is whether the taxi was yellow. The number of observations is based on all registered taxis (ranging from 15,588 to 16,933 per month) in the 36-mo period. R^2 is the coefficient of determination. ***$P < 0.0001$, **$P < 0.001$.

Dataset S1. Accident replication code

Dataset S1

Dataset S2. Thirty-six-month accident record

Dataset S2

Dataset S3. Thirty-month, 3,341-driver accident and demographic record (the 3-mo, 3,341-driver daily driving record is in three separate files)

Dataset S3

Dataset S4. Distribution of daily driving time in different districts

Dataset S4

Dataset S5. Distribution of daily driving time in different statuses

Dataset S5

Dataset S6. Average speeds of yellow and blue for the 168 h of the week

Dataset S6