Boise State University

ScholarWorks

9-2019

Misalignment Between Societal Well-Being and Business Profit Maximization: The Case of New York Taxis Drivers' Incentive System

Andrew S. Manikas
University of Louisville

James R. Kroes
Boise State University

Thomas F. Gattiker
Boise State University

Misalignment Between Societal Well-Being and Business Profit Maximization: The Case of New York Taxis Drivers' Incentive System

Andrew S. Manikas*
University of Louisville
Management Department
College of Business
Louisville, KY
andrew.manikas@louisville.edu

James R. Kroes
Boise State University
College of Business and Economics
Boise, ID

Thomas F. Gattiker
Boise State University
College of Business and Economics
Boise, ID

Abstract

Objectives of business sustainability efforts commonly include increasing consumer safety, decreasing resource consumption, and decreasing pollution. Even though there is a societal interest in attaining these goals, business and other economic agents often operate under incentive structures that run counter to these objectives. Taxi drivers operate as economic independents. Their revenue depends on their fares and tips. Moreover they choose how many hours to work, how fast to drive, and which route to take. Using New York City taxi data from 2013, we test the level of alignment between the revenue maximizing behavior of drivers versus safety, conservation and pollution- related outcomes that are valued by stakeholders. We find substantial misalignment—i.e., in order to maximize revenue, drivers take inefficient routes and they exceed the speed limit thus decreasing safety, increasing fuel consumption and increasing air pollution. Based on these empirical results, we suggest methods of aligning societal goals with those of revenue maximizing taxi drivers.

Keywords: digital data; taxi; New York City; policy alignment; public safety

1. Introduction

Sustainability objectives at the business and societal levels often include increasing consumer safety, decreasing resource consumption, and decreasing pollution. However, business and other economic agents often operate under incentive structures that run counter to these objectives. This forces trade-offs – often undesirable trade-offs – between economic sustainability and social/environmental sustainability.

Taxi drivers are rational, profit seeking economic actors. They have a great deal of discretion when it comes to decisions such as how long to work, what route to use for each fare, and how fast to drive. These decisions affect each driver's profits. These decisions also affect each driver's sustainability footprint – including factors such as safety risk, resource consumption, pollution generation, and other social and environmental outcomes. We use data collected from the New York City taxi fleet in order to explore the level of alignment between drivers' profit outcomes and their sustainability-related outcome. As we report below, we find substantial misalignment between these outcomes – i.e., in order to maximize revenue, drivers take inefficient routes and often exceed the speed limit thus decreasing safety, while increasing fuel consumption and air pollution. Based on these empirical results, we suggest methods of aligning societal goals with those of revenue maximizing taxi drivers.

2. Literature Review

Economists have frequently examined if individuals are rational and profit maximizing. In their seminal paper on New York taxi driver behavior, Camerer et al. (1997) identified that taxi drivers operate as independent economic workers. Drivers select how long to drive, and they keep the fares and tips as their revenue. The driver behavior should give good insights into rational behavior. Camerer et al. (1997) suggested also that drivers have a set point of money they intend to make each day. Once they reach that set point, the driver will stop working. This would thus explain why on rainy days, when demand for their services is high, it is hard for customers to find an available cab. Rather than working more hours on days where demand is high and extra money could be made easily, drivers stop working when they reach their revenue set point. This work was refuted by Farber (2005) who found that drivers typically stopped when they got tired, usually after 12 hours on the shift. Additionally, Farber (2008) concluded that the stopping condition for a driver's shift was strongly related to hours worked, and if a set point of revenue earned was related to driver probability of stopping, that set point changed each day. The idea that taxi driver stopping criteria was based on hours worked rather than revenue earned was confirmed in Crawford and Meng (2011). Farber (2015) found that drivers tend to learn how to make more revenue per hour as they gain experience driving over a number of years. It seems more likely that since the number of cabs allowed in New York is limited (via medallions), the increased demand for taxis in the rain is unmet because the taxi capacity is not flexible. In addition, too many taxis would make the career unappealing as during normal weather and times, it would be extremely difficult for each driver to make a living. Because each driver generally works until they are tired, revenue earned is reflective of a profit maximizing economic individual. By focusing on revenue (fare and tip) per fare, the number of hours worked before the driver quits for the day should not affect our study results here. Therefore, our sample frame is the individual fares (trips) rather than the entire shift as done in prior research.

In New York City, taxi drivers pay a flat fee to use the cab, and then keep the revenue and tips. All else being equal, higher tip percent equates to more revenue for the driver for a particular fare. In their study on tipping, Haggag and Paci (2014) utilized a payment screen that suggested default tips (in various amounts). They found that setting default tips (even though a button for "other amount" was available) made people tip less. It appears that customers do not like the suggestion of a default and will tip based on other factors, including the perception that the driver is meeting their transportation needs. In New York City, taxi meters determine the fare through a combination of time and distance measures. The standard city rate (Rate Code 1) charges customers a $2.50 flat fee upon entry as well as a $0.50 charge per additional unit. A unit is defined as either (i) a 60 second interval in which the car is idle or driving less than 6 miles per hour or (ii) 0.20 miles when the car is driving 6 miles per hour or faster. Additionally, there are surcharges, which are captured in the dataset; specifically; 1) a night surcharge of $.50 between 8:00 PM and 6:00 AM, 2) a peak hour weekday surcharge of $1.00 Monday through Friday between 4:00 PM and 8:00 PM, and 3) a New York State tax surcharge of $.50 per ride (NYC.gov A, 2017).

Given this fare calculation methodology, taxi drivers may possibly have an incentive to exceed the speed limit or take a longer route between two locations. For example, as shown in Figure 1 (an example trip from the Empire State Building to the American Museum of Natural History), a driver has a number of options available at his discretion on how to complete the trip. The shortest distance by road is 3.0 miles (Route 1) - assuming the trip is driven at 18 miles per hour (MPH) and also assuming for this simplified example there is no idle time at intersections, it takes 10 minutes (via 5th avenue, cutting through Central Park at 79th Street Traverse). Similarly, this trip could take 11 minutes to go 3.0 miles (Route 2) by cutting through Central Park at 65th Street Traverse. The fare to the driver would be the same, but by choosing Route 1 the driver can make that amount ($2.50 flat fee + [0.50 * 5] * 3 miles = $10) in 10 minutes versus 11 minutes. This would be optimal for the driver and the passengers. However, going out of the way (Route 3, i.e. inefficient routing), the trip can become 3.2 miles, adding 50 cents to the driver fare. The driver would now have to go 19.2 MPH to complete this slightly longer trip (9th Ave to Columbus Ave) in the same 10 minutes. Alternately, the driver could drive at 18 MPH on this longer mileage route and take a little less than 11 minutes to arrive at the destination. If the passengers were not familiar with the streets or traffic in the city, 10 minutes and 11 minutes may both seem appropriate for their trip. However, the longer route is not beneficial to the passengers (unless a significant traffic event was blocking the lower mileage routes).

© OpenStreetMap contributors

Figure 1: Taxi driver route options from the Empire State Building (marked A bottom of figure) to the American Museum of Natural History (marked B at top of figure).

The speed limit in New York since 1986 has been 30 MPH, but in 2014 was lowered to 25 MPH in an effort to reduce traffic deaths (Bankoff, 2014). Speeding cars are more likely to cause fatalities as the risk of death is approximately four times higher when a pedestrian is hit at 40 MPH than at 30 MPH (Department for Transport, 2014). In fact, speeding killed 9,262 people in 2014 (NHTSA, 2015). Speeding also causes injuries and has a negative economic impact on society. The 2016 Taxi & Limousine Commission (TLC) Factbook (NYC.gov D, 2016) notes that beyond the death toll, the total economic impact of speeding averages about $40.4 billion a year in the United States. Further, the TLC Factbook details that over 99% of the taxi drivers are male. This compounds the likelihood of speeding as a 2015 government study (NYC.gov E, 2017) noted that men were more likely to speed than women (66 percent versus 61 percent). Taxi drivers seem to have incentive to speed, while society wants them to drive more carefully. Our analysis will look at how taxi drivers can maximize their revenue, and maximize tip percent, on fares. We assume taxi drivers to be rational agents (as done in prior economics studies using taxi drivers), thus the drivers should evince behavior consistent with earning the most revenue for their work effort.

3. Research Model and Empirical Analysis

Donovan and Work (2014) obtained data from the New York City Taxi and Limousine Commission through a Freedom of Information Law (FOIL) request. The dataset covered four years, the most recent being 2013. We use this last year of data for our study here. Because the medallion and hack licenses are reassigned each year, it is only possible to track drivers (hack licenses) and vehicles (medallions) within a calendar year.

3.1 Hypotheses

Beyond 12 MPH speed, a driver makes more money per time period driving versus sitting idle with a passenger because the miles driven at the fare rate of 50 cents per 0.2 miles contribute to revenue more than tips – reported as 18% on average (NYC.gov C, 2014). Thus, a rationally economic driver would want to drive faster to make more revenue per hour. Only if the trip has an average speed of less than 12 MPH will the trip time contribute more to the fare than the miles driven. Therefore, a driver who makes one trip of five miles can make more money with five trips of one mile each, as each new trip has a $2.50 entry fee. This equates to the driver making an additional $10 ($2.50 x 4) all else being equal for having five one mile trips compared to a single five mile fare. However, having a fare paying passenger in the taxi seems preferable to dropping them off and looking for another passenger and then looking for more customers three more times to each pay the $2.50 entry fee. It seems better to have a guaranteed passenger generating revenue than risk driving empty or sitting idle while waiting for the next possible fare. Therefore, we expect to see that driving faster is associated with more revenue for a driver:

H1a: Average driver speed is positively associated with drivers' revenue per hour.

Assuming that drivers are economic actors who operate to maximize their individual economic profit, we expect drivers to augment their revenue by driving faster; however, customers do not share this incentive. Moreover customers value safety - passengers generally associate speed with recklessness and they may worry about injuries to themselves, others, or damaging inanimate objects in a crash when in a speeding taxi. We do recognize that when people are running late, they may prioritize speed over safety, but presumably most passengers are not running late most of the time. Therefore, we expect passengers to express their desire for safety by tipping a smaller percent of the fare when the taxi driver goes faster:

H1b: Average driver speed is negatively related to drivers' tip percentages.

As suggested by H1a, drivers that can drive more miles for a fare can make more revenue. We define inefficiency (INEFF) as the trip distance (recorded on the meter) divided by the straight line air distance (e.g., flying from point A to point B, without having to turn on streets). A score of one would be a perfectly efficient route, and scores greater than one indicate increasingly circuitous routing. The implication is that driving less efficiently (higher INEFF) by taking a fare on a longer route increases revenue because it increases miles for which a fare is charged. This is a well-documented way for drivers to "rip off" customers (Deaton, 2011). Peer and Solomon (2012) found that professional drivers as well as non-professional drivers had considerable errors in their estimations of driving speeds and journey times. However, the researchers determined that overestimations of time for taxi drivers were smaller than overestimations by nonprofessional drivers. Thus passengers may overestimate the distance or time to get to their destination, allowing a savvy cab driver to take a less efficient route adding miles and fare revenue to the trip total. Assuming that drivers will not be caught and sanctioned, taking a route that is longer than the optimal route is a way for drivers to maximize their individual economic well-being. Thus we expect to see:

H2a: Driver inefficiency (INEFF) is positively associated with revenue

Conversely, passengers have a preference for the most efficient route and they might be aware that a less efficient route is being used. Some passengers know the city and traffic conditions as well or nearly as well as taxi drivers, and thus they know when the driver is taking a less than efficient route. Even less knowledgeable riders may sense that the driver is not taking them the most direct route. For example, the driver may make many seemingly unnecessary left and right turns. Moreover, most passengers have GPS enabled mapping application on their phones. This allows them to draw some conclusions about route efficiency—for example, by comparing the actual route to the route that approximates a straight line, or by asking the app to map out the route. Finally, some apps estimate travel time from origin to destination, and passengers can compare this time to the actual travel time. In these cases, passenger's conclusions about whether the driver is taking the most efficient route are imperfect, but they provide some indication that the driver is deviating from the best route. Since customers prefer that drivers take the best route, we expect that:

H2b: Driver inefficiency (INEFF) is negatively associated with tip percent

3.2 Model

We use multiple regression, as done in McLay et al. (2012), to test our hypotheses. Looking at which factors were significant in predicting responses via multiple regression was also done by Iacocca et al. (2013).

3.3 Data

Consistent with other studies (e.g. Bornstein et al., 2013), to make the data set manageable, we use a randomly selected sample of 1% of drivers over all 12 months of 2013 (the most current year of NY Taxi data available at the commencement of this study). Then those hack licenses (uniquely associated with drivers) were used to filter out all other drivers from the 2013 data. The tip amount was rarely recorded when the payment type was cash. Therefore, we use only credit payment trips recorded in the dataset. Our final sample consist of 949 drivers making 262,224 trips. To include possible weather effects, we retrieved the hourly weather for New York City (weathersource.com, 2016). We assume that the weather recorded at the observatory in Central Park on a particular day and hour would be similar for all of the city. The variables of interest in the data set are shown in Table 1:

Table 1: Variable descriptions

Variable Name	Description
Revenue Efficiency (REVEFF)	Fare amount plus tip amount per second of the trip
Tip percent (TIPPCT)	Tip Percent (TIPPCT) = tip amount / fare amount
Speed (SPEED)	Calculated as trip distance / trip time in seconds (Miles Per Second), then converted to miles per hour (MPH)
Rain (RAIN)	Set to 1 for light, moderate, or heavy precipitation, 0 for no precipitation
Vendor (VENDOR)	Vendor of the electronic payment machine in cab (shows fares/tips, takes credit cards, etc.): 1 = *VeriFone Transportation Systems* (VTS), or 0 = *Mobile Knowledge Systems Inc.* (CMT)
Holiday	Set to 1 for major holidays in the USA, 0 otherwise
Weekend	Set to 1 if the day of week is Saturday or Sunday for the trip, 0 otherwise
Passenger count (PASSCT)	The number of passengers recorded on the fare
Driver Inefficiency (INEFFF)	Trip distance / straight line air distance (flying from point A to point B, without having to turn on streets). A score of one would be a perfectly efficient route, and scores greater than one indicate increasingly circuitous routing.

Data Filtering:

We used similar data filters as Haggag and Paci (2014) to minimize data entry/capture errors:

- We removed toll rides and flat fare trips (e.g., to/from airports).

- We dropped duplicate fares where all variables were the same or the same medallion and trip start were identical, since that is impossible.

- We dropped fares where payment type was "No Charge" or "Dispute"

- We removed cash transactions because payments of cash almost all have no tip amount recorded by the driver. Credit card payments auto collect the tip amount. 45% of entries were cash and were dropped as tip amounts were needed for this study. As a note, prior research has looked at anchoring to raise tip amounts on credit card transactions (Haggag and Paci, 2017). Further, at least one news article stated that drivers get bigger tips when people use credit cards, perhaps because of the suggested amounts in the Haggag and Paci (2017) system. However, we would suggest a possible sinister explanation; namely, taxi drivers may underreport cash tips in an effort to avoid paying mandatory Federal income, Social Security, and Medicare taxes (IRS, 2016). Given that over 99% of the cash transactions in our study had a zero tip while the average credit card tip was approximately 17.6%, it seems possible that the difference is a recording error by the

driver rather than a distinctly different group of passengers that use cash versus credit card. It seems that the government could require taxi drivers to pay tax on tips for cash fares at the same percentage as the automatically recorded tips on credit card fares.

- We dropped fares where the trip time in seconds was zero as that must be a data error.

- We dropped fares where the trip distance was zero as that must be a data error.

- We eliminated fares if the tip amount was negative, which must be a data error.

- We dropped fares where the latitude or longitude were impossible values (i.e., outside of the possible ranges -90/90 and -180/180, respectively).

- Any fare greater than the maximum calculated fare given distance and time was removed from the sample. A maximum fare was calculated based on all possible surcharges plus the entire trip time as idle time plus the entire mileage as units.

The main variables used in the analyses were winsorized at the 1% level to further remove any outliers.

Dependent Variables:

Our analyses test two basic models, each with unique dependent variables:

- The first model uses Revenue Efficiency (REVEFF) as the dependent variable. REVEFF, calculated as the fare amount plus tip amount per second of the trip, measures how effective a driver is at generating revenue during a trip.

- Tip Percent (TIPPCT) is the dependent variable in the second model. TIPPCT, measured as the tip amount / fare amount, gauges the additional value the customer felt they received during a trip.

Independent Variables:

To test our hypotheses, we employ the following two independent variables:

- Trip speed (SPEED) is the average speed of a trip in miles per hour (MPH).

- Driver Inefficiency (INEFF) equals the trip distance / straight line air distance (flying from point A to point B, without having to turn on streets). A score of one would be a perfectly efficient route, and scores greater than one indicate increasingly circuitous routing. When selecting the use of the straight line air distance in our measure, we assumed that, over a sample as large as the one in our study (n=262,224), on average the air distance would be proportional to the most direct driving distance - particularly since our sample was drawn from trips in Manhattan, which has a street system largely organized around a rectangular grid plan. To justify this use of straight line air distance as a proxy for driving distance, we collected actual shortest path driving distance data to compare with the straight line distance data used in our measure. To do this, we used a Google Maps API that allows users to query actual driving distance data using a MS Excel macro (the actual driving distance from Google Maps does consider one-way streets and other restrictions.) Unfortunately, collecting full data for our sample proved impractical as the API limits the number of queries to 2,500 per day. Therefore, we collected actual driving distance data for 5% of the trips used in our analysis, which were randomly selected from our sample (i.e. 13,112 observations). We then tested the correlation between the straight-line distances used in our study with the actual driving distances retrieved from Google Maps. We found that these two measure have a correlation of 0.92, significant at the $p<0.0001$ level. We believe that this high level of correlation supports the use of the straight line air distance as a proxy for the shortest driving distance in our inefficiency measure.

Controls:

The following variables are included in our models to control for other factors that might impact the revenue efficiency and tip percentage:

- *WEEKEND*. Set to 1 for Saturday and Sunday, 0 otherwise. Weekend travelers may differ from week day taxi cab users who may be more likely to be business people going to and from work.

- *HOLIDAY*. Set to 1 for major US holidays, 0 otherwise. Passenger traffic and customer demand on holidays might differ from normal patterns.

- Passenger count (PASSCT). The number of passengers does not affect the fare based on miles or time, but may take longer to load and unload, thus reducing the number of fares a driver can take per hour. Further, tips may be higher if passengers want to impress co-riders, or lower if people feel that other passengers should chip in more.

- *VENDOR*. Two vendors supply the in taxi payment systems; 1) VeriFone Transportation System (VTS), and 2) Mobile Knowledge Systems Inc. (CMT). Note that the CMT system is easier to fix if a passenger wants to pay by credit card, but the driver already selected cash. As noted by the New York City TLC (NYC.gov B, 2017), on the CMT system, there is a back button on the Driver Information Monitor (DIM), while on the VTS system, the driver needs to enter a negotiated fare (Rate Code 5), and then allow the passenger to pay by credit card. This minor difference may alter tips as a customer might be annoyed by the complexity on the VTS system. A dummy VENDOR was created with 0 for CMT and 1 for VTS.

- Rain (RAIN). We expect that precipitation might impact the behaviors of both drivers and passengers. To control for this possible effect, precipitation in inches was obtained from the Central Park observatory hourly data (weathersource, 2016). RAIN is set to 1 if there is precipitation or snowfall, otherwise it is 0 if no precipitation is present.

Model Specification:

Our main models to test, given that *i* is a specific fare, are:

$$REVEFF_i = \beta_0 + \beta_1(WEEKEND_i) + \beta_2(HOLIDAY_i) + \beta_3(PASSCT_i) + \beta_4(VENDOR_i) + \beta_5(SPEED_i) + \beta_6(RAIN_i) + \beta_7(INEFF_i) + \varepsilon_i$$

$$TIPPCT_i = \beta_0 + \beta_1(WEEKEND_i) + \beta_2(HOLIDAY_i) + \beta_3(PASSCT_i) + \beta_4(VENDOR_i) + \beta_5(SPEED_i) + \beta_6(RAIN_i) + \beta_7(INEFF_i) + \varepsilon_i$$

Table 2 lists the minimum, maximum, and mean values for some of the variables of interest. The mean speed of 17 MPH, calculated as the trip_distance / trip_time_in_sec (miles per second) and then converted to miles per hour (MPH), is well below the 30 MPH speed limit present during 2013. Though the average speed is below the legal limit, of the 262,224 trips in the sample, 11,976 (4.6%) had an average speed in excess of 30 MPH. Further, 221 of the 969 (22.8%) drivers in the dataset had at least one fare during the year with an average speed above 30 MPH. Unfortunately, an exact determination of the number of trips during which a driver exceeded the speed limit is impossible as the available data does not contain idle time, only total miles driven and total duration of the trip. Hence, while the exact number cannot be calculated, the actual number of trips in which the driver exceeded 30 MPH for a portion of the trip is certainly larger because the average speed calculation does not account for stops and idle time. Combined, these data points suggest that speeding is in fact a real problem for New York taxi drivers and passengers.

Table 2: *Variable Statistics*

Variable Name	Min	Max	Mean
REVEFF (per second)	0.01	0.08	0.02
TIPPCT	0.00%	60.00%	17.59%
SPEED (MPH)	7.76	36.25	17.62
Average Trip Time (Minutes)	9.00	33.00	18.57
Average Trip Distance (Miles)	2.47	11.51	5.44

4. Results

Our multiple regression results are shown in Table 3. Models (1) and (3) are the base models used to test our hypotheses as specified in the in the prior section. Models (2) and (4), which are discussed in the next section, conduct post-hoc tests of interaction effects of interest. Models (1) and (2) have an outcome variable of revenue efficiency (*REVEFF*), and models (3) and (4) have an outcome variable of tip percent (*TIPPCT*). All of the models have significant F-Statistic values, indicating that each of the regression models significantly predict the respective outcome variables. To insure that multicollinearity among our sample is not influencing the results, we calculated the variance inflation factor (VIF) score for each of the variables and found that the VIF scores range between 1.0 and 1.09, well below the recommended threshold of 10.

From Model 1, we see that H1a and H1b are both supported ($p<0.01$); drivers make more revenue by speeding and passengers tend to give lower tip percentages for speeding, thus the faster speed allows a driver to collect revenue more quickly even though the tip portion of revenue may be lessened. The results of our second analysis (Model 3) indicate support for both H2a and H2b – demonstrating that drivers make more revenue by not taking the most efficient route with their fares and that passengers seem aware that they are being driven "the long way" and subsequently lower their tip percent.

Table 3: *Regression results with outcomes of Revenue Efficiency (REVEFF) and Tip Percent (TIPPCT)*

VARIABLES	(1) REVEFF	(2) REVEFF	(3) TIPPCT	(4) TIPPCT
WEEKEND	6.61e-05***	6.68e-05***	-0.00297***	-0.00297***
	(1.13e-05)	(1.13e-05)	(0.000282)	(0.000282)
HOLIDAY	0.000148***	0.000147***	-0.000199	-0.000197
	(2.42e-05)	(2.41e-05)	(0.000604)	(0.000604)
PASSCT	-2.58e-05***	-2.58e-05***	0.000220**	0.000220**
	(3.95e-06)	(3.94e-06)	(9.87e-05)	(9.87e-05)
VENDOR	-0.000155***	-0.000154***	-0.00460***	-0.00460***
	(1.10e-05)	(1.09e-05)	(0.000274)	(0.000274)
SPEED	0.000698*** **(H1a)**	0.000699***	-0.000552*** **(H1b)**	-0.000553***
	(8.75e-07)	(8.75e-07)	(2.19e-05)	(2.19e-05)
RAIN	-0.000154***	-0.000156***	-9.42e-05	-8.29e-05
	(1.42e-05)	(1.42e-05)	(0.000355)	(0.000355)
INEFF	0.000200*** **(H2a)**	0.000375***	-0.000380*** **(H2b)**	-0.000676***
	(2.55e-06)	(8.26e-06)	(6.37e-05)	(0.000207)
SPEED x INEFF		-6.07e-06***		1.07e-05
		(2.68e-07)		(6.71e-06)
RAIN x INEFF		5.40e-05***		-0.000272
		(7.49e-06)		(0.000187)
Constant	0.00859***	0.00859***	0.189***	0.189***
	(1.85e-05)	(1.84e-05)	(0.000461)	(0.000461)
Observations	262,224	262,224	262,224	262,224
F-Statistic	92,712***	72,319***	155.6***	121.5***
R-squared	0.712	0.713	0.004	0.004

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Note the high R-squared value for the outcome variable revenue efficiency (Models 1 and 2) in Table 3. This intimates that a driver could reasonably predict their approximate revenue per hour based on what day of the week it was, if it was a holiday, what electronic payment (*VENDOR*) they had installed, the weather (*RAIN*), and under their control, how fast they drove and how inefficiently they travelled with customers. Conversely, tip percentage models (3 and 4) have substantially lower R-squared values, which suggests that other unmeasured factors not included in our model might impact the tip percentage. This may be related to socio-economic factors of the passengers that are not available in the recorded fare data.

It is possible that the fixed fees of fares could increase the revenue efficiency per our hypothetical example in Section 3.3 where a driver with five one mile trips makes more than a driver making a single five mile trip because each fare pays a fixed fee to enter the taxi. We ran a post hoc analysis with REVEFF as the outcome where we isolated short (5 miles or less) and long (greater than 5 miles) trips by creating a new dummy variable SHORT as shown in Table 4.

Table 4: *Regression results with outcome of Revenue Efficiency (REVEFF) for the full model from Table 3, but with addition of a dummy variable SHORT (1: <=5 miles, 0: >5 miles fare distance)*

VARIABLES	(1) REVEFF
WEEKEND	-7.12e-05***
	(1.08e-05)
HOLIDAY	2.23e-05
	(2.30e-05)
PASSCT	-2.08e-05***
	(3.76e-06)
VENDOR	-0.000150***
	(1.04e-05)
SPEED	0.000794***
	(1.02e-06)
RAIN	-0.000100***
	(1.35e-05)
INEFF	0.000375***
	(7.88e-06)
SHORT	0.00198***
	(1.23e-05)
SPEED x INEFF	-6.14e-06***
	(2.56e-07)
RAIN x INEFF	5.28e-05***
	(7.15e-06)
Constant	0.00577***
	(2.48e-05)
Observations	262,224
R-squared	0.739

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

As Table 4 shows, the revenue efficiency is enhanced by the driver taking shorter trips. This makes sense because each new fare pays a fixed fee to enter the taxi. As the example noted in Section 3.3, a driver could make an additional $10 ($2.50 x 4) all else being equal for having five one mile trips compared to a single five mile trip. If the driver can keep the taxi full, multiple short distance fares could increase their revenue efficiency.

5. Discussion

This paper examines drivers' economic incentives versus customers' interests in efficiency and safety. We find that drivers have an incentive (i.e. higher fares) to drive long distances and to do so at high speeds. Passenger tipping behavior indicates that passengers do express a desire for efficiency and safety. However, passengers' tips do not provide a sufficiently strong disincentive to prevent drivers from taking longer than optimal routes and doing so at higher speeds.

As an example, if the most efficient route for a trip has a distance of 5 miles, at 30 MPH that takes 10 minutes to complete with no stops. The fare would be $2.50 flat fee plus 5 miles * (5 units * $0.50) for a total fee of $15. Assume the happy passengers tip 18%, for a total revenue of $17.70. Conversely, an unethical driver could take them on a 7 mile trip to the same destination by going out of the way. If the driver sped at 42 MPH, the trip would still be 10 minutes with no stops. However, now the fare would be a $2.50 flat fee plus 7 miles * (5 units * $0.50) = $20.00. Even if the passengers felt the drive was circuitous and did not leave a tip, this trip would still generate more revenue for the taxi driver compared to the efficient route example ($20.00 versus $17.70).

In addition to the interests of drivers and passengers, there is a societal interest in safety (for pedestrians, cyclists, and property, as well as for motorists and their passengers) as well a societal interest in resource efficiency—especially fuel conservation and pollution prevention. Unfortunately, customer tipping behavior along with speed limits and other regulations is not enough to align taxi drivers' behavior with these broader interests. An important part of the incentive structure is the fare calculation, instituted by New York City, which allows drivers to make the most revenue in a set period of time by speeding and taking inefficient routes. Regulatory measures might improve the situation. However, a proposal in 2014 by the Mayor of New York to shut off meters when the driver is speeding was not well received by taxi drivers and has not been implemented (Fasick et al., 2014). Another component of this Mayoral proposal was to require more training for cab drivers who crash frequently.

In examining the results, the presence of rain, which was included as a control, does have an effect on the revenue efficiency for the taxi driver. Looking at fares with precipitation versus without, we calculated a mean trip speed of 17.0 MPH with precipitation versus 17.7 MPH without precipitation. This lowering of speed is likely the cause of the decreased revenue. However, our data does not show passengers to be altering tip percentage based on the presence of precipitation. They may be happy to be riding in a taxi versus walking in the rain, but their tips do not reflect this.

To investigate this issue in more depth, we conducted a post-hoc analysis of driver behavior to see which variables affect their speed (see Table 5). This showed that drivers tend to drive faster on weekends and holidays. Also of interest, drivers speed up as they become more inefficient in their route. This makes intuitive sense as shown in the example in Figure 1. A route out of the way (more inefficient) gives more revenue to the driver, but the passengers are less likely to notice the extra mileage if the trip takes approximately the same time as the direct route. Thus, a revenue maximizing strategy is for a taxi driver to drive inefficiently and quickly. However, the low R-squared value in Table 5 suggests that other factors may contribute to the decision to drive faster.

Table 5: *Post Hoc Speed Analysis*

VARIABLES	(1) Speed
WEEKEND	0.234***
	(0.0252)
HOLIDAY	1.195***
	(0.0538)
PASSCT	-0.000635
	(0.00881)
VENDOR	0.0384
	(0.0244)
RAIN	-0.442***
	(0.0316)
INEFF	0.0494***
	(0.00569)
Constant	17.52***
	(0.0229)
Observations	262,224
F-Statistic	152.0***
R-squared	0.003

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

As an additional post-hoc analysis, we examined if the combination of driving an inefficient route while speeding or when it is raining impacts revenue and tips. To test this, we modified our original multiple regression models to include interactions between drive inefficiency with both speed and rain. The first analysis, Model 3 in Table 3, shows a significant interaction effect of speed and inefficiency on revenue efficiency. The negative coefficient indicates that at a certain very high level of speed and inefficiency, revenue efficiency would start to decline. However, as presented in Figure 2 (which shows the interaction effect for a speed range of plus and minus one standard deviation around the

mean speed), the interaction coefficient is not very substantial despite its statistical significance - which means that a driver will make more money by driving faster and less efficiently except in the most extreme of cases. Rain and inefficiency also had a significant, but small interaction effect on revenue efficiency. Figure 3 shows the differences when there was rain during the fare, and where there was no rain. Again, being more inefficient is best for the driver's revenue per hour. Whether it is raining does have a statistically significant effect on revenue, but practically speaking it is a small effect. Regardless of the weather, drivers make more money by taking the longest route they can get away with in most circumstances.
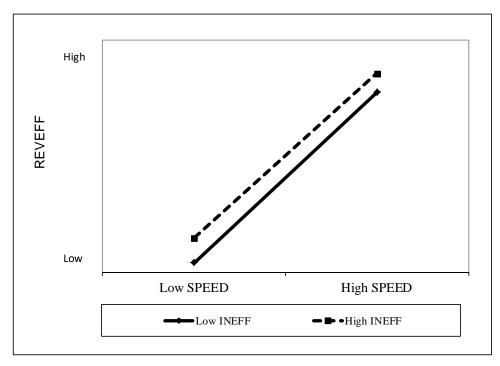

Figure 2: *Plot of interaction between speed and inefficiency. More inefficient routes result in higher revenue efficiency (REVEFF) regardless of speed. Low and High are calculated as mean +/- 1 std. dev for each of the measures. For example, the x-axis covers speeds of 11.7 – 23.5 MPH.*
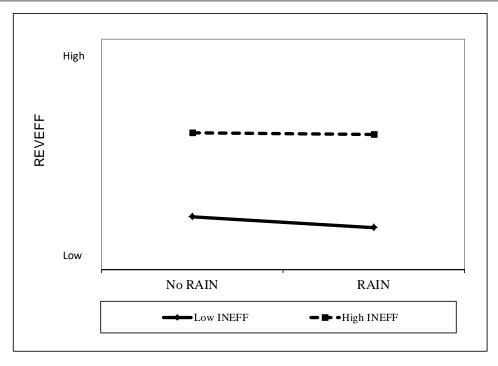
Figure 3: *Plot of interaction between rain and inefficiency. Higher inefficiency results in more revenue efficiency for the driver. Drivers taking efficient routes, have lower revenue efficiency in high rain conditions (likely due to the reduced speed they end up driving in the rain). Low and High are calculated as mean +/- 1 std. dev for each of the measures.*

6. Limitations

A limitation of our study is that the data is from New York City taxi fares in the year 2013 only. We did not use data from other cities, but we would expect that similar reward schemes would induce drivers in other cities and other years to behave similarly.

A further limitation of the available data is that we cannot know the actual most efficient route for each fare. Roads blocked and other factors cause the best route to change over time. As a proxy for the most efficient route we used the point to point distance (as the crow flies). This would only be the actual most efficient route for a trip that went along a straight line only. The most efficient route for most fares would be longer than the point to point distance. However, all routes in our analysis were calculated with this proxy for efficient route so we believe that the relative distance deviation from efficiency is still valuable in lending insights.

7. Conclusions and Recommendations

In summary, our analysis clearly shows that taxi drivers have in incentive to act against the interests of passengers and of society and that taxi drivers do indeed behave in accordance with these incentives. We provide recommendations that would change the incentive structure to align it with societal and customer interests.

Although the original proposal in 2014 to shut off meters when the driver was speeding was not implemented, we believe this would be a positive step. The current speed limit in New York City is now 25 MPG, thus any taxi driver exceeding that rate would be identified automatically by the electronic fare system. Sending them speeding tickets would not be as effective as the immediate feedback of the meter turning off (and possibly some visual signal such as a warning light or sound). Customers would be alerted to the excessive speed and the driver would not only be forfeiting fare during the speeding incidents, but perhaps the tip might also decrease as passengers would be aware that the driver was acting illegally and dangerously. This loss of fare and tip revenue would better align the goals of the taxi driver with those of the city and passengers. The public's voice must be heard along with taxi drivers' – surveys can find what community members think and what policies make sense. Ernst and Lambe (1983) designed and implemented a survey of people using parking meters. A similar survey instrument could be developed to

determine what the public's feelings are on speeding taxis. If the community at large responds that safety is more important than other factors, then taxi drivers wishing to maximize revenue by speeding might be discouraged to speed with new fare/meter rules (e.g., shutting the meter off if taxi is speeding).

The issue of taxi driver taking inefficient routes could be mitigated. The optimal driving distance in miles could be calculated via a routing software and GPS technology when the fare ends (the pickup and drop off latitude and longitude are recorded by the taxi system automatically now). Perhaps the fare according to mileage could be automatically lowered to a threshold. For example, if the miles driven were more than 10% over the calculated optimal route, the fare based on mileage would be set to the minimum of the actual or 10% over. Idle time is not in the driver's interest and often outside their control (waiting at traffic lights), so fee units would still be added (for when the vehicle was below 6 MPG). Kourtit et al. (2016) proposed the use of digital data systems for smart city policy. With the existing taxi data used for this study and the real time in car tracking systems, better decisions can be made on public transportation needs, and the state of taxi driver safety practices.

References

Bankoff, C. New York Daily Intelligencer, November 7, 2014. http://nymag.com/daily/intelligencer/2014/11/things-to-know-about-nycs-new-speed-limit.html

Bornstein, M. H., Jaegar, J, and Putnick , D. L.. Sampling in Developmental Science: Situations, Shortcomings, Solutions, and Standards. PubMed 2013; 33(4): 357-370.

Camerer, C. Babcock, L., Lowenstein, G., and Thaler, R. Labor Supply of New York City Cab Drivers: One Day at a Time, The Quarterly Journal of Economics 1997; 112(2): 407-441.

Crawford, V. P., and Meng, J. "New York City Cab Drivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income†, American Economic Review 2011; 101: 1912–1932.

Deaton, J. P. "How Taxi Meters Work", 31 March 2011. HowStuffWorks.com. http://auto.howstuffworks.com/taxi-meter.htm (accessed June 18, 2017)

Department for Transport. THINK. http://think.direct.gov.uk/speed.html (accessed June 17, 2014).

Donovan, B., and Work, D. New York City Taxi Trip Data (2010-2013), University of Illinois at Urbana-Champaign. Dataset. http://dx.doi.org/10.13012/J8PN93H8, 2014.

Ernst, F. F., and Lambe, T. A. Design and Implementation of a Survey of People Parking at Curbside Meters. Socio-Economic Planning Sciences 1983; 17(3): 121-129.

Farber, H. Is Tomorrow Another Day? The Labor Supply of New York City Cabdrivers, Journal of Political Economy 2005; 113(1): 46-82.

Farber, H. Reference-Dependent Preferences and Labor Supply: The Case of New York City Taxi Drivers, American Economic Review 2008; 98(3): 1069–1082.

Farber, H. Why You Can't Find a Taxi in the Rain and Other Labor Supply Lessons from Cab Drivers*, The Quarterly Journal of Economics 2015; 130(4): 1975-2026.

Fasick, K., Fonrogue, G., and Harshbarger, R. De Blasio wants to shut down meters to slow speeding cabs. New York Post, February 18, 2014. http://nypost.com/2014/02/18/de-blasio-wants-to-shut-down-meters-in-speeding-cabs (accessed June 15, 2017)

Haggag, K., McManus, B., and Paci, G. Learning by Driving: Productivity Improvements by New York City Taxi Drivers†, American Economic Journal: Applied Economics 2017; 9(1): 70–95.

Haggag, K., and Paci, G. Default Tips, American Economic Journal: Applied Economics 2014; 6(3): 1-19.

Iacocca, K., Sawhill, J., and Zhao, Y. A multiple regression model to explain the cost of brand-drugs, Socio-Economic Planning Sciences 2013; 47: 238-246.

IRS. Tips on Tips – A Guide to Tax Income Reporting (for Employees who Receive Tip Income). United Stated Department of the Internal Revenue Service. https://www.irs.gov/pub/irs-pdf/p3148.pdf (accessed 2/24, 2018).

Kourtit, K., Nijkamp, P., and Steenbruggen, J. The significance of digital data systems for smart city policy. Scoio-Economic Planning Sciences 2016, http://dx.doi.org/10.1016/j.seps.2016.10.001

McLay, L. A., Boone, E. L., and Brooks, J. P. Analyzing the volume and nature of emergency medical calls during severe weather events using regression methodologies. Socio-Economic Planning Sciences 2012; 46: 55-66.

NHTSA. National Highway Traffic Safety Administration – United States Department of Transportation 2015. https://www.nhtsa.gov/risky-driving/speeding (accessed 6/17, 2017)

NYC.gov A. NYC Taxi & Limousine Commission.
http://www.nyc.gov/html/tlc/html/industry/taxicab_rate_yellow.shtml (accessed June 13, 2017)
NYC.gov B. NYC Taxi & Limousine Commission. 2017.
http://www.nyc.gov/html/tlc/html/industry/taxicab_serv_enh_archive.shtml (retrieved June 14, 2017)
NYC.gov C. NYC Taxi & Limousine Commission. 2014 TLC Factbook.
NYC.gov D. NYC Taxi & Limousine Commission. 2016 TLC Factbook.
NYC.gov E. NYC Health. https://www1.nyc.gov/site/doh/about/press/pr2017/pr016-17.page (accessed June 17, 2017)
Peer, E., and Solomon, L. Professionally biased: misestimations of driving speed, journey time and time-savings among taxi and car drivers, Judgment and Decision Making 2012; 7(2): 165-172.
Weathersource. https://weathersource.com/ (accessed 12/17/2016)