# Cheating When in The Hole: The Case of New York City Taxis

Shivaram Rajgopal\* Schaefer Chaired Professor in Accounting Goizueta Business School Emory University shivaram.rajgopal@emory.edu

and

Roger White PhD Student, Accounting Goizueta Business School Emory University roger.white@emory.edu

June 22, 2015

Preliminary and Incomplete Comments welcome

#### **Abstract:**

Fraud occurs when there is an incentive and the opportunity to commit fraud and the fraudster can rationalize his behavior. Academic literature in financial economics has investigated incentives and opportunities to commit fraud but has largely ignored the fraudsters' ability to rationalize fraud. We address this gap by positing that economic actors, whose earnings are restricted by regulations, cheat more and rationalize such behavior as an effort to get "out of the hole" that regulations have forced them into. In a sample drawn from 200 million New York City (NYC) taxi rides for 2013, we find that taxi drivers are unlikely to overcharge (0.5% likelihood) customers when dropping them off in areas in which NYC taxis are legally able to collect another passenger, but are very likely to overcharge passengers (80%-90% likelihood) when dropping passengers off in areas in which NYC taxis are not allowed to pick up another customer. We interpret these findings as evidence that regulatory restrictions on earnings can prompt economic actors to view themselves as "in the hole," which leads to (much) higher levels of cheating and fraud.

\*Contact author

We thank Olivia Burnsed, Ilia Dichev, Ron Harris, Lisa Laviers, Grace Pownall, Jesse Schwartz, Jared Smith, and Han Stice for feedback that greatly improved this manuscript. We also thank numerous cab drivers who gave freely of their time and good nature in answering questions from two naïve accountants. All errors are our own.

# Cheating When in The Hole: The Case of New York City Taxis 1.0 Introduction

Why do people commit fraud? Statement on Auditing Standards SAS No. 99 states that three conditions are generally present when fraud occurs: (1) the incentive to commit fraud, (2) the opportunity to commit fraud, and (3) the ability to rationalize fraud. There is considerable work on the incentives to commit fraud (e.g., Dechow, Sloan, and Sweeney (1996), Erickson, Hanlon, and Maydew (2006), Lie (2005), Bergstresser and Philippon (2006), Burns and Kedia (2006), Efendi, Srivastava, and Swanson (2007), Armstrong et al. (2013)) and some on the opportunity to commit fraud (in particular, poor governance as in Beasley (1996) and Farber (2005)). However, there is virtually no large sample research on the fraudster's ability to rationalize fraud partly because it is hard for an empiricist to observe the fraudster's micro day to day behavior. In this study, we examine the rationalization of fraud, in particular, whether regulations make it easier for individuals to rationalize fraud. We are able to do so because we have access to micro data on 200 million cab rides taken during the year 2013 in New York City.

In particular, some rides take cab drivers outside of the city limits of New York City (NYC), where NYC cabbies are prohibited from picking up new customers. Compared to similar trips in which cabbies drop off passengers within the city limits of New York, we find fraud to be much more likely for these out-of-city destinations. Hence, in a setting where regulations governing a transaction restrict earnings in the future, service providers are more likely to defraud the current customer in an effort to make up for the reduced future earnings. That is, individuals seem willing to cheat in an effort to make sure they do not end up "in the hole."

In some respects, our research design mimics a regression discontinuity approach. We examine a sample of very similar trips from Brooklyn (primarily) to Staten Island or New Jersey. The Staten Island destinations are just as distant as the New Jersey destinations, and the rides last just as long. The only difference is that the New Jersey rides entail crossing a border, a border that delineates where NYC cabbies can pick up new customers. We find that crossing this border into New Jersey, which means the cabbie has to backtrack into NYC to pick up a new fare, greatly increases the likelihood that the cabbie overcharges the passenger. This regulation indirectly restricts future earnings, as driving back to NYC before finding another fare understandably takes time and fuel.

This restriction increases the likelihood of fraud by a factor of at least 50. The Brooklynto-Staten Island trips see cabbies overcharge their passengers about 1% of the time. In a very similar sample of New Jersey bound trips, we observe overcharging on 50% to 90% of rides, depending on the destination. It appears that cabbies are very likely to commit fraud in an effort to avoid ending up "in the hole" as a result of regulatory restrictions on where they can do business.

This result improves our understanding of where and when fraud occurs. That is, fraud is situational in nature. In robustness tests we control for cab driver fixed effects, and our observed effect persists at comparable levels. This suggests that the same cabbies who are unlikely to defraud customers on trips within the city are very likely to fraudulently overcharge customers when they are placed in situations that likely lessen future earnings (ferrying customers to New Jersey).

Our finding informs research on predictors of fraud, both individual (personal) and situational. Individual level covariates linked to fraud prediction include testosterone (increases

fraud likelihood, Jia, Lent, and Zeng (2014)), military service (lowers fraud likelihood,

Benmelech and Frydman (2014), Koch and Wernicke (2013)), and tenure, age, and education (see Zahra, Priem, and Rasheed (2005) for a review). Importantly, none of these covariates have much predictive power. For example, Benmelech and Frydman (2014) find that managers with military experience commit fraud about half as often as their non-military counterparts, and Jia, Lent, and Zeng (2014) find that high testosterone managers are about twice as likely to commit fraud as low testosterone managers. While these changes are statistically significant, they still lead to pseudo R-squared measures of less than 0.25 in binary models predicting fraud occurrence. By comparison, New Jersey bound rides in our sample (that come with decreased expected future earnings, as cabbies must back track into New York City to pick up another fare) see fraud at least 50 times (5,000%) higher than comparable rides in New York City. Pseudo R-squared statistics suggest that at least 80% of the variation in fraud is predicted in these models. While this comparison of pseudo R-squared measures is informal, it provides some degree of the importance of situational pressures in predicting fraud.

All told, this evidence is strongly suggestive of situational influences playing a large role in inducing and predicting fraud. Burns and Kedia (2006), Bergstresser and Philippon (2006), and Efendi, Srivastava, and Swanson (2007), among others, put forth evidence that managers are more likely to commit fraud when option incentives are on the line. Our result suggests that this inducement to commit fraud is driven not only by utility for wealth, but by a behavioral bias to avoid ending up "in the hole" relative to expected earnings.

We expand on this idea in the following section, where we outline more specific theory and articulate our hypothesis. Further sections describe the data, empirical tests, and results, as well as a series of robustness tests, followed by conclusions.

#### 2.0 Theory and Hypothesis Development

#### 2.1 Factors affecting fraud

Statement on Auditing Standards, SAS 99, states that three conditions are generally present for fraud to occur. First, there must be incentives or pressure to commit fraud. Second, an opportunity for committing fraud must present itself. Finally, there must exist an attitude or rationalization for committing fraud.

Of these three drivers of fraud, incentives is perhaps the most heavily researched area. For instance, Dechow, Sloan, and Sweeney (1996) find that an important motivator for accounting fraud is to attract external financing at a lower cost. Erickson, Hanlon, and Maydew (2006) find no association between an executive's equity incentives and accounting fraud. However, Efendi, Srivastava, and Swanson (2007) find a link between "in the money" options and earnings misstatements. Burns and Kedia (2006) find the delta of an executive's stock option is associated with misreporting whereas Armstrong et al. (2013) seem to attribute that result to the vega of the executive's stock option portfolio. Lie (2005) suggests that hundreds of firms might have intentionally backdated their stock options.

Turning to opportunities, several papers have linked poor governance to fraud. Dechow, Sloan, and Sweeney (1996) find that firms that manipulate earnings are less likely to have independent boards, an audit committee, and an outside blockholder, and more likely to have the CEO as the chairman of the board. Beasley (1996) also links a lack of board independence to fraud. Farber (2005) finds similar results but also finds that fraud firms are less likely to have financial experts on their board and are less likely to be audited by Big 4 audit firms.

Attitude or rationalization to commit fraud has received the least attention in the literature. In a field survey, Nelson, Elliott, and Tarpley (2002) find that managers are more

likely to structure transactions to achieve certain reporting outcomes when the accounting standards that govern them are excessively precise. Agoglia, Doupnik, and Tsakumis (2011) find that CFOs in their laboratory experiment are less likely to report aggressively when they encounter principles-based, as opposed to rules-based, accounting standards. It is difficult to document large sample evidence on the importance of rationalizations or attitudes towards fraud partly because it is hard for a researcher to obtain data on a fraudster's attitudes. Alternatively, the researcher needs access to micro activities the fraudster indulges in on a day to day basis to infer such attitudes or rationalizations. We are able to access such micro data for 200 million cab rides in NYC for the 2013. In addition to bringing such data to bear on the discussion, we propose the use of the hyper-motivation as the psychological force driving much of the rationalization towards cheating in our setting. We turn to a detailed discussion of that force next.

#### 2.2 Rationalization

Rick and Loewenstein (2008) observe that much dishonest behavior stems from individuals that view themselves as "in the hole." Dishonesty, in the minds of these individuals, is a means to "catch up" to what they view is an appropriate economic position. Rick and Loewenstein (2008) define motivations of this type as *hyper-motivation*. While Rick and Loewenstein (2008) were the first to formally define hyper-motivation, other behavioral scholars have recently made similar observations. Notably, Cameron, Miller, and Monin (2010) observe that individuals in a loss frame (versus an economically equivalent profit frame) are more likely to upwardly misreport their performance on an experimental task, where such misreporting leads to increased pay. Hyper-motivation is a construct closely related to loss aversion (Kahneman and Tversky (1979)). Cameron and Miller (2009) includes an insightful discussion of this relation, which centers on the idea that a loss frame increases the acceptance of one's own unethical behavior (Brewer and Kramer (1986)). Consequently, cheating in an effort to avoid realizing such a loss (or minimize the loss) is more likely.

We examine the importance of hyper-motivation in predicting fraud in a real world setting. While there is much research on the predictors and covariates of fraud, most is either (1) experimental or (2) focuses only on observed (typically prosecuted/disciplined) fraud. Both of these literatures are valuable and informative, but both suffer from weaknesses. Specifically, the experimental literature is unable to address the economic magnitudes of different predictors of fraud outside of the laboratory, as experiments typically only vary two or three such predictors. Likewise, field research is typically limited to addressing only instances in which fraud is actually observed. Understandably, this is less than complete sample of total fraud, and is biased towards those who are unsuccessful at keeping fraud hidden (Zitzewitz (2012)).

The literature in forensic economics generally avoids these issues by examining data that allows for the observation or estimation of actual (underlying) instances of fraud (see Zitzewitz (2012) for an excellent literature review). One famous example of this approach is Jacob and Levitt (2003), who examine standardized test scores in Chicago schools to determine which teachers exhibit evidence of tampering with exams (changing students answers). Consistent with hyper-motivation, teachers in schools on probation (for prior poor test performance) were more likely to cheat.

Similar evidence exists in financial markets. For example, Burgstahler and Dichev (1997) find that firms are much more likely to report barely positive (or zero) earnings per share

than negative earnings per share. Work by Grundfest and Malenko (2009) and Bhojraj et al. (2009) finds that these firms (that barely beat expectations) are more likely to restate earnings per share downwards and experience negative long run returns. Similar to underperforming teachers (Jacob and Levitt (2003)), one could argue that managers who view themselves as "in the hole" are much more likely to commit fraud in an effort to break even (or come out ahead). Likewise, evidence from tax research suggests that individual taxpayers are more likely to cheat on their taxes when they expect to owe money to tax authorities, as opposed to when they expect to receive an economically equivalent refund (Chang and Schultz (1990), Schepanski and Kelsey (1990)).

Jacob and Levitt (2003) and Burgstahler and Dichev (1997), as well as the underlying theory in Rick and Loewenstein (2008), deal generally with instances in which prior poor performance puts economic actors "in the hole." Being "in the hole" then leads to hypermotivated cheating. We test whether regulation can have similar effects. For example, Morgenstern, Pizer, and Shih (2001) estimate that complying with environmental regulations increases costs across steel, plastic, paper, and petroleum plants by about 2%. If plant managers view their pre-regulatory costs and expected earnings as a reference point, then after the costs of environmental compliance, they may view themselves as "in the hole." If this is the case, then hyper-motivation may lead to increased fraud and cheating among populations where profits are restricted by regulation.

Testing this hypothesis, stated in the alternate form below, is the goal of this paper.  $H_A$ : Regulatory restrictions on profits lead economic actors to engage in more frequent cheating behavior.

#### 3.0 Methods and Data

To test our hypothesis, we use data from the New York City Taxi and Limousine Commission's (TLC) Taxi Passenger Enhancement Program (TPEP). The TLC collects TPEP data, via a GPS system built in to taximeters in NYC cabs, on every ride given by licensed taxis. This data includes date, time, distance traveled, time elapsed, origin, destination (both via latitude and longitude), fare, tolls, tip, method of payment, cab driven, and cab driver (anonymized license ID). The TPEP system works by recording the GPS coordinates and time of day whenever a driver activates the meter to begin a fare. At this stage, the driver also inputs the number of passengers. The taximeter then records the distance and fare for the trip as it takes place. At the conclusion of the trip, indicated by the driver turning off the taximeter, the TPEP system transmits the final GPS coordinates, starting GPS coordinates, time, fare, and other ride characteristics (date, driver ID, cab ID, etc.) to the TLC data servers.

Chris Whong, a civics enthusiast, procured this data (for year 2013) from the TLC via a Freedom of Information Act (FOIA) request. Mr. Whong subsequently posted the data, with some accompanying analysis, on his website for public consumption.

Similar data is used by Haggag and Paci (2014), who examine tip levels as a function of suggested tips, Farber (2014), in an examination of labor supply, Haggag, McManus, and Paci (2014), in a study of learning, and King and Peters (2012), who research route choices as a function of distance, time, and cost. Labor economists have also made use of taxi trip sheets, the analog version of this data, in multiple studies (Farber (2008), Camerer et al. (1997), Ashenfelter, Doran, and Schaller (2010)).

Our particular setting examines fraud committed by taxi drivers at toll booths. Specifically, cab drivers are required to use the E-Z Pass lanes. These lanes are typically much faster than cash lanes, as drivers do not have to stop, pay tolls via cash, and receive change. Instead, drivers with an E-Z Pass receiver can drive through specially designated E-Z Pass lanes without stopping. E-Z Pass readers built into the toll booth read the driver's receiver as the car passes (at speed), and deducts the toll directly from a bank account the driver has registered to their E-Z Pass receiver.

In addition to allowing for quicker passage, E-Z Pass lanes are also about 25% cheaper, on average, than cash lanes for the same toll booth. Regulatory authorities set up such differential pricing to encourage E-Z Pass adoption among drivers in an effort to reduce traffic congestion around toll booths.

Taxi drivers in New York City are required by the TLC to use E-Z Pass lanes, and TLC penalties for using cash lanes instead of the E-Z Pass range from \$150 to \$200 (NYC TLC (2010)). However, taxi drivers have economic incentives to use cash lanes for two reasons. First, cash lanes are notably slower than E-Z Pass lanes, and driver compensation (taxi fare) is a function of both time and distance, so slower rides lead to higher driver pay. Second, cash lanes are more expensive. While the difference in total costs is paid to the toll booth, the passenger's tip is typically a function of total costs. Accordingly, increasing total costs by going through a cash toll lane typically results in the driver receiving a higher tip. This benefits taxi drivers directly, as most cabbies in NYC rent a taxi by the week (per shift) and are the residual claimants of any revenue.

Many customers are unaware of the requirement that NYC cabs use E-Z Pass lanes, and a recent investigative report by the *New York Post* discovered that cabbies frequently take advantage of this ignorance and use cash lanes (Namako (2010)). The TPEP data allows for such fraudulent overcharging to be detected perfectly, as E-Z Pass tolls are completely

distinguishable from cash tolls by amount.<sup>1</sup> For example, cash tolls tend to be round numbers, whereas E-Z Pass tolls are not. Using a pricing schedule for tolls from New York's Metropolitan Transportation Authority, we can determine whether each toll reported in TPEP data is an amount that corresponds to an MTA cash toll amount or an MTA E-Z Pass toll amount. There are no overlapping amounts (amounts that could be for either a cash or E-Z Pass toll).

The total TPEP data set posted by Mr. Whong contains nearly 200 million taxi rides for the entirety of 2013. We examine a small minority of these rides, specifically those that cross the Verrazano-Narrows Bridge that connects Brooklyn to Staten Island.<sup>2</sup> See Figure 1 for a map of the area.

The Verrazano-Narrows Bridge is unique in that the cash toll is \$15, while the E-Z Pass toll is only \$10.66. This is the largest absolute discount (and also the highest tolls) offered by the E-Z Pass system, and presumably offers the greatest incentive for drivers to upcharge their customers. Such effects may be undetectable or absent in the case of the more typical \$7.50 cash toll that is discounted to \$5.33 for E-Z Pass users (e.g., Queens–Midtown Tunnel, Robert F. Kennedy Bridge).

Furthermore, this directed analysis of the Verrazano-Narrows Bridge allows us to test whether hyper-motivation can be triggered by regulatory restraints. For much of Brooklyn, the quickest way to the Newark Liberty International Airport in New Jersey is via Staten Island and the Verrazano-Narrows Bridge. New York City cabs are required to take passengers to Newark Liberty upon request, but cannot pick up a passenger on the return trip until crossing back into

<sup>&</sup>lt;sup>1</sup> Despite this ability, enforcement by the TLC is still done manually on the spot. That is, when the TLC is concerned that taxi drivers are using cash lanes at toll booths, instead of analyzing the TPEP data and sending tickets to offenders, the TLC sends staff to go stand in cash toll booths (with the toll attendant) and hand out tickets directly to taxis passing through (Namako (2010)).

 $<sup>^{2}</sup>$  We also restrict our analysis to rides that take place after March 2, 2013, as toll prices were increased permanently on that date, and limiting our tests to a single regime of toll prices reduces noise and allows for cleaner identification of a treatment effect.

New York City. That is, New York City cabs are required to ferry passengers to an out-of-state airport, but cannot offer their services for a return trip to customers interested in getting to New York City from Newark Liberty (Haddon (2012)).

If cab drivers view themselves as "in the hole" when picking up a fare destined for Newark Liberty (as they cannot carry a paying passenger on the return trip), then hypermotivation may lead them to be more likely to defraud these customers (use the cash lane on the Verrazano-Narrows Bridge) in an effort to raise their earnings up to what they view as a more appropriate (or fair) level.

Teasing out this effect necessitates controlling for other characteristics of the taxi ride that may influence a cab driver's likelihood of fraudulently using the cash lane on the Verrazano-Narrows Bridge (VanderWeele and Shpitser (2011)). Accordingly, we include taxi ride characteristics (fare, time, distance, method of payment, passengers), driver income (dummies for highest and lowest tercile of earnings in the previous month), and time characteristics (dummies for rush hours, winter months, and weekdays). To proxy for congested traffic, we also include a variable that measures the number of NYC cabs to pass through the Verrazano-Narrows Bridge (bound for Staten Island) in a given hour (we refer to this covariate as *traffic*). We winsorize the extreme 1% of values of our continuous variables to reduce the effect of statistical outliers.

#### 4.0 Results

Before explicitly testing our hypothesis, we estimate a series of models to confirm that taking the cash lane results in a longer trip and a larger tip. We focus exclusively on trips to Staten Island (excluding trips to New Jersey) in this preliminary analysis, as it results in a more homogeneous sample.

Panel A of Table 1 reports the summary statistics for data used in this preliminary analysis. Notably, only about 1% of taxi rides heading from New York's other Boroughs (primarily Brooklyn<sup>3</sup>) to Staten Island involve a taxi driver defrauding their passenger by using a cash lane on the Verrazano-Narrows Bridge. On the whole, these trips are typically long (median of 35 minutes, 16 miles) and relatively expensive (median fare of \$46, not counting toll and tip).

Panel B of Table 1 reports the cross-sectional regressions predicting time and tip as a function of the cabbie using the cash lane, in addition to other controls. Model 1 estimates the time each trip takes, controlling (notably) for distance, traffic, and time of day. The variable of interest (*Cash Lane at Toll*) is a dummy variable that equals one if the cabbie used a cash lane (as opposed to E-Z Pass, as required by law) when crossing the bridge. As expected, the cash lane takes longer. The significant (p=0.08) coefficient of 103.0 suggests that taking the cash lane adds about a minute and 43 seconds to the cab ride (103 seconds). NYC cabs charge \$0.40 per minute idle (stuck in traffic), which would suggest that the fraudulent tactic of using the cash lane increases the driver's expected fare by about \$0.69.

Models 2 and 3 estimate similar regressions using the driver's tip as the dependent variable. As the cash toll (\$15) is higher than the E-Z Pass toll (\$10.66), the base cost of a trip using the cash lane is higher by about \$4. If customers tip as a function of this base cost, then drivers fraudulently using the cash lane may see an increased tip. Model 2 uses all fares between NYC's other four boroughs and Staten Island, whereas Model 3 uses only fares paid by credit

<sup>&</sup>lt;sup>3</sup> Most of these rides originate in Brooklyn, as trips to Staten Island and Newark Liberty from Manhattan, The Bronx, and much of Queens are faster by way of the Lincoln Tunnel, Holland Tunnel, or George Washington Bridge.

card (for which tips are reported perfectly).<sup>4</sup> The significant (p<0.05), positive coefficient on the cash lane dummy in both models suggests that tips are about \$1 higher when drivers use the cash lane.

These preliminary tests suggest that, all told, cab drivers very rarely defraud customers by using the cash lane for rides to Staten Island from the other boroughs. When cabbies do engage in this (rare) fraudulent behavior, it increases their earnings for the trip by about \$2. Note that in addition to the increased costs paid to the cabbie, customers victimized by this fraudulent behavior are also responsible for paying the cash toll that is about \$4 higher than the E-Z Pass rate.

As a final preliminary analysis, we model cabbies' use of the cash lane using a logistic regression.<sup>5</sup> Table 2 reports these results. Intuitively, and in line with previous research (e.g., Porcano (1988), Dimmock and Gerken (2012)), a powerful predictor of contemporaneous fraud is past fraud. Cabbies who have previously used the cash lane on the Verrazano-Narrows Bridge are more than ten times more likely to defraud current customers. Specifically, with other controls set at the median, cabbies who have no history of defrauding customers (in our sample) by using the cash lane on the Verrazano-Narrows Bridge are only 0.5% likely to do so on the current trip. By comparison, cabbies who have previously defrauded customers in this manner (about 3% of rides) do so on future rides (other controls set at median) about 5.5% of the time.

<sup>&</sup>lt;sup>4</sup> The TPEP system does not record cash tips for taxi rides. Farber (2014) and Haggag and Paci (2014) mention this, and we confirm it via correspondence with Joanne Rausen, the TLC's Assistant Commissioner for Data and Technology.

<sup>&</sup>lt;sup>5</sup> King and Zeng (2001) find that conventional logit models are poorly specified when there is little variation in dependent variable outcomes (<5% of dependent variables differ from the mode). This is the case in our data, where fraud only occurs about 1% of the time for rides ending in Staten Island and originating in other NYC boroughs (data for Table 1 and Table 2). These authors (and Tomz, King, and Zeng (2003)) construct a rare events logit model well defined for such uncommon events, and in untabulated analysis we confirm that our Table 2 results are consistent in this alternative specification.

Only two other notable significant predictors of fraud emerge from this model, method of payment and a driver having low income. Cabbies are less likely to defraud customers who pay with cash. Method of payment is unobservable when the fraud decision is made, but perhaps cabbies can predict method of payment and prefer cash for tax avoidance reasons. Alternatively, customers paying in cash are likely to be poorer, as credit card access skews towards wealthy individuals (Cohen-Cole (2011)). Prior work has found that service providers are less likely to cheat poorer individuals, which may be driving this negative coefficient on cash payment (Gino and Pierce (2010)). Finally, the fact that lower income cabbies are more likely to commit fraud is consistent with hyper-motivation, and generally is in line with prior laboratory experiments (John, Loewenstein, and Rick (2014)), as well as archival research on tax fraud (e.g., Slemrod, Blumenthal, and Christian (2001), Christian (1994), Orviska and Hudson (2003)).

All told, this preliminary analysis of New York City taxi rides (crossing the Verrazano-Narrows Bridge) suggests that fraud is relatively rare (occurs less than 1% of the time). In general, taxi drivers tend to take the E-Z Pass lane when crossing the bridge, even though doing so typically results in a lower tip and lower fare than the fraudulent alternative of using the cash lane. Also, customers seem susceptible to this type of fraud, and regulatory enforcement is lax (Namako (2010)).

In the next set of tests, we examine whether a regulatory quirk, the prohibition on NYC cabs picking up fares in New Jersey, leads to a higher level of fraud in the trips bound for the Newark Liberty Airport. The TPEP data we use contains about 5,500 such trips, see Panel A of Table 3. Notably, these trips look very similar in most respects to typical trips that end on Staten Island (Table 1 and Table 2). For example, the median trip to the Newark Airport (Staten Island) lasts 34 (35) minutes and covers 17.9 (15.7) miles. Fares are noticeably different, but mostly

because the trip to the Newark Airport increases the base cost of a NYC taxi trip by \$15 (per mile and per minute costs remain the same, but drivers add a \$15 surcharge directly to Newark Liberty bound trips).

This increased fare is intended to offset the cost to drivers of not being able to pick up another customer until they return to New York City. If this measure is effective, then cab drivers will likely not feel hyper-motivated pressure to cheat, as they will not view themselves as "in the hole." However, this does not appear to be the case. Summary statistics in Panel A (Table 3) suggest that a full 92% of fares to the Newark Airport that cross the Verrazano-Narrows Bridge do so using the cash lane as opposed to the E-Z Pass lane (as mandated by regulators). We model this propensity in a logistic regression in Panel B.

Panel B of Table 3 reports a logit model estimating the likelihood that taxi drivers (fraudulently) use the cash lane when carrying customers across the Verrazano-Narrows Bridge to Staten Island (30,341 observations) or the Newark Airport (5,473 observations). The coefficient on the Newark Airport dummy is positive and significant (p<0.0001), indicating that cabbies are much more likely to use the cash lane when ferrying passengers to the Newark Airport. In terms of economic significance, with other controls set at the median, a driver is only 0.4% likely to use the cash lane when taking a passenger to a Staten Island destination, but 89.4% likely to use the cash lane when taking a passenger to the Newark Airport.

This result, indicative of the vast majority of taxi drivers fraudulently overcharging customers headed to Newark Liberty, is very much in line with hyper-motivation and our stated hypothesis. Drivers appear to view themselves as "in the hole" on such trips, as picking up a return fare is prohibited by regulatory restrictions. Consistent with Rick and Loewenstein (2008), Cameron and Miller (2009), and Cameron, Miller, and Monin (2010), this mindset

appears to make cheating easier to sanction psychologically, which leads to increased cheating and fraud. The degree of this increase is particularly surprising. Drivers who showed almost zero fraudulent behavior (< 1%) in typical rides within New York City are almost all fraudsters (~ 90%) when conducting an otherwise synonymous transaction with regulatory restrictions on future earning ability (by prohibiting a paying passenger on the return trip).

#### **Robustness Tests**

While we attribute the Table 3 results to hyper-motivation, an alternate explanation is that NYC cabbies are much more likely to defraud Newark Liberty bound customers (by using the cash lane when crossing the Verrazano-Narrows Bridge) because these riders are more often tourists (or other travelers) who are unlikely to know that such actions are prohibited. Balafoutas et al. (2013), for example, find in a field experiment that customers who claim not to be familiar with a city are more likely to be taken on long, costly detours by cabbies. This is perhaps not a pressing concern in our setting, as the NYC residents interviewed in a newspaper investigative report on toll overcharging were not familiar with the regulation either (Namako (2010)), but for completeness we attempt to rule out this potential explanation.

To do so, we include in the data set an addition 662 taxi rides that involve passengers crossing the Verrazano-Narrows Bridge on trips bound for New Jersey to points other than the Newark Liberty Airport. As these trips are not destined for the airport, they are more likely made by area residents who are more familiar with local taxi regulations. As familiarity with regulations is less a concern, the observed cheating behavior observed by taxi drivers in this subsample is more likely attributable to hyper-motivation.

Additionally, this subsample contains transactions that are perhaps more generalizable. Fares to Staten Island and the Newark Airport are set by regulators, but fares to other points in

New Jersey are completely negotiable, and cabbies have the option to refuse service. That is, cabbies only accept fares to New Jersey locations excluding Newark Liberty on their own terms, which maps better into more common transactions that are not typically regulated in terms of price.

Panel A of Table 4 reports the summary statistics of these rides, and we add them to the sample of within-NYC rides (Panel A, Table 1) and NYC-to-Newark Airport rides (Panel A, Table 3). Panel B reports the results of a logit model, similar to those in Table 2 and Panel B of Table 3, which predicts the likelihood that a cabbie uses the cash lane when crossing the Verrazano-Narrows Bridge. The variable of interest in this model is the indicator for rides that end in New Jersey at locations other than the Newark Liberty Airport, *Destination: NJ (non-Newark Airport)*.

This variable of interest loads with a significant (p<0.0001) positive coefficient, indicating that rides to New Jersey (excluding the Newark Airport), like rides to the Newark Airport, are very likely to see cabbies use the cash lane when crossing the Verrazano-Narrows Bridge. Even for rides in which cab drivers set the terms, hyper-motivation induced by being unable to pick up a return fare appears to induce cheating behavior. This variable is powerful in terms of economic significance. With controls set at the median, the likelihood of a driver using the cash lane (fraudulently) on a Staten Island bound trip is 0.4%. Shifting the dummy variable of interest, *Destination: NJ (non-Newark Airport)*, to one increases this expected likelihood to 72.3%. This result strongly suggests that hyper-motivation, rather than taking advantage of tourists or a reaction against a strongly regulated pricing scheme, is responsible for the fraudulent overcharging (via cash lanes on the Verrazano-Narrows Bridge) we observe. Next, we examine linear probability models that predict the likelihood of a cabbie using the cash lane. The primary advantage of linear probability models in our setting is that they allow for the inclusion of driver fixed effects, even when drivers only enter the sample once. A logistic regression with fixed effects drops such observations (and all such cases in which a driver fixed effect perfectly predicts cash lane usage). While logit models are much more commonly applied to dichotomous dependent variables, linear probability models are generally well specified in large samples using heteroskedastic robust standard errors (see White (1980), Petersen (2009)) when the predicted probabilities are largely contained between zero and one (see Hellevik (2007), Angrist and Pischke (2009)).

Table 5 reports the linear probability models predicting cabbies using the cash lane in our complete sample (intra-city rides, Newark Airport rides, and other NJ rides). Model 1 replicates the baseline model reported in Table 4, Panel B, and Model 2 includes driver fixed effects. Notably, the coefficients of interest (*Destination: Newark Airport* and *Destination: NJ (non-Newark Airport)*) are similar and highly significant across specifications. *Destination: Newark Airport* has a coefficient around 0.80, which suggests that trips to the Newark Airport are about 80% more likely than the baseline to see the cabbie use the cash lane on the Verrazano-Narrows Bridge. The corresponding coefficient for *Destination: NJ (non-Newark Airport)* is about 0.50. While these effect sizes are smaller than those obtained in the logit models, they still indicate a huge increase in the likelihood of cheating behavior among cabbies in transactions that place regulatory restrictions on future earning potential. For example, with controls set at the median, the predicted likelihood of a cabbie using the cash lane for an intra-city trip (ending in Staten Island) is about 0.8% (Model 1, Table 5). For trips ending in New Jersey (excluding the Newark Airport), this corresponding estimated likelihood is about 52.5%.

Most importantly, that our results persist when controlling for driver fixed effects suggests that individual driver characteristics are not responsible for our findings. Rather, controlling for unobservable individual characteristics (e.g., honesty, aggression, empathy), hyper-motivation appears to still induce cheating behavior in rides (to New Jersey) that cabbies view as putting them "in the hole."

As a final robustness test, we investigate whether or not our results are driven by cabbies that may not have access to a functioning E-Z Pass. In response to the *New York Post* report on cabbies using cash lanes, at least one prominent taxi blogger suggested that such activity stems from cabbies that do not use E-Z Pass at all (or very rarely).<sup>6</sup> This response centered on the fact that most cabbies are immigrants, who are more likely to be unbanked (Rhine, Greene, and Toussaint-Comeau (2006), Rhine and Greene (2006)). Unbanked individuals are, by definition, unlikely to use an E-Z Pass, as an E-Z Pass must be affiliated with a bank account (to bill for toll charges). While this does not explain why we would observe differential rates of overcharging for trips bound for New Jersey, it is nonetheless a potentially important covariate to control for. We do so partially in our primary analyses with the *Prior Cash Lane at Toll* measure, but that only captures past use of cash lanes at the Verrazano-Narrows Bridge toll booth.

For a better control, we build a measure of a driver's propensity to use an E-Z Pass overall, *Driver Used E-Z Pass for* >50% of tolls in prior month. This dummy variable equals one for all drivers who used an E-Z Pass for the majority of their toll transactions, across all toll booths in New York City, in the past month. Notably, this control variable equals one in over 99% of the driver-months in our sample,<sup>7</sup> suggesting that the attribution in the taxi blogosphere is probably incorrect. It does not appear that unbanked cabbies (without E-Z Passes) are driving

<sup>&</sup>lt;sup>6</sup> See http://yellowtaxiblog.com/reasons-behind-a-taxi-drivers-using-cash-lane-in-bridges-and-tunnels/

<sup>&</sup>lt;sup>7</sup> In more than 90% of our driver-months, E-Z Pass usage at toll booths exceeds 99%. That is, the vast majority of drivers use E-Z Pass for almost all of their toll payments.

our results. To be certain, we estimate our primary logit models in Table 6 including the aforementioned dummy variable (*Driver Used E-Z Pass for >50% of tolls in prior month*). Model 1 includes only rides to Staten Island that cross the Verrazano-Narrows Bridge, Model 2 includes rides to Staten Island and Newark Liberty, and Model 3 includes trips to Staten Island and all of New Jersey (including Newark Liberty).

*Driver Used E-Z Pass for* >50% *of tolls in prior month* loads as negative and significant, suggesting that the few taxi drivers that frequently (and fraudulently) use cash lanes at most toll booths (and are likely to be unbanked or liquidity constrained to the point of not keeping much money in their bank account) are more likely to (fraudulently) use the cash lane at the Verrazano-Narrows Bridge. However, this variable moderates our hyper-motivation findings only by a small margin. Consider the case in Model 2, with all controls set to the median, even for drivers who use E-Z Pass for the majority of their toll booth charges, shifting the destination from Staten Island to Newark Liberty increases the likelihood of the cabbie fraudulently using the cash lane from 0.4% to 88.2%. The corresponding increase for drivers who did not use E-Z Pass for the majority of fraudulent use cash lanes over the Verrazano-Narrows Bridge in general (higher base rate), but still see hyper-motivation increase cheating by a great deal when regulations restrict (future) earnings (i.e. for trips to Newark Liberty). Results are similar for Model 3, which includes trips to other New Jersey destinations.

#### **5.0 Conclusions**

Rick and Loewenstein (2008), Cameron and Miller (2009), and Cameron, Miller, and Monin (2010) put forth the idea, and some experimental evidence, that individuals who view themselves as "in the hole" are more likely to commit fraud in an effort to improve their

financial standing (hyper-motivation). We examine whether regulatory restrictions that indirectly curb earnings can incite this hyper-motivated reasoning. In a sample of taxi rides in New York City, we find that cabbies rarely overcharge customers (< 1%) when dropping them off in areas where NYC taxis are legally able to collect another fare (in the five boroughs). However, on trips ending in New Jersey, where NYC taxis are not permitted to pick up new passengers, almost all cabbies overcharge (~ 90%) their customers. These rides are otherwise very similar (in terms of distance, time, time of day, and number of passengers), which suggests that it is this indirect regulatory restriction on earnings, and cabbies' subsequent self-evaluation of being "in the hole", that drives increased cheating/fraud.

Regulations play an important role in developed, Western-style economies. Modern, well-regulated economies, for example, largely avoid issues like milk tainted with poisonous chemicals (as in China in 2014, Branigan (2009)) and rivers so polluted that they catch on fire (as in Ohio in 1952 and 1969 and China in 2014, see Adler (2014), Malm (2014)). However, our results suggest that regulations that restrict earning potential (like using safe processing practices for milk or properly disposing of industrial pollutants) increase fraud, which should be considered by regulators of all stripes.

More importantly, our results offer some idea of the magnitude with which hypermotivation can induce fraud. Identifying potential fraudsters is difficult (Albrecht et al. (2011), Kranacher, Riley, and Wells (2010)), and research on personal characteristics related to fraud has only identified relatively weak individual level predictors (e.g. Jia, Lent, and Zeng (2014)). By comparison, our work suggests that situational characteristics, rather than personal characteristics, may offer helpful insight to the auditor, manager, or customer looking to detect or avoid fraud. For example, Jia, Lent, and Zeng (2014) find that high testosterone CEOs are twice as likely to misreport financial performance. By contrast, in our setting, cabbies who view themselves as "in the hole" (due to regulatory restrictions) are more than 50 times more likely to overcharge passengers. While these settings are very different, the comparison is at least indicative of the powerful psychological motivations at play in individuals who view themselves as "in the hole." We believe we are the first to put forth large sample evidence on how fraudsters rationalize fraud, and given the increased predictive power controlling for such rationalization offers, we encourage future researchers to incorporate similar controls, when possible.

### References

Adler, Jonathan H., 2014, The fable of the burning river, 45 years later, The Washington Post.

- Agoglia, Christopher P., Timothy S. Doupnik, and George T. Tsakumis, 2011, Principles-Based versus Rules-Based Accounting Standards: The Influence of Standard Precision and Audit Committee Strength on Financial Reporting Decisions, *The Accounting Review* 86, 747–767.
- Albrecht, W. Steve, Chad O. Albrecht, Conan C. Albrecht, and Mark F. Zimbelman, 2011, *Fraud Examination*. 4th edition. (Cengage Learning, Mason, OH).
- Angrist, Joshua D., and Jörn-Steffen Pischke, 2009, *Mostly Harmless Econometrics: An Empiricist's Companion* (Princeton University Press, Princeton, NJ).
- Armstrong, Christopher S., David F. Larcker, Gaizka Ormazabal, and Daniel J. Taylor, 2013, The relation between equity incentives and misreporting: The role of risk-taking incentives, *Journal of Financial Economics* 109, 327–350.
- Ashenfelter, Orley, Kirk Doran, and Bruce Schaller, 2010, A Shred of Credible Evidence on the Long-run Elasticity of Labour Supply, *Economica* 77, 637–650.
- Balafoutas, Loukas, Adrian Beck, Rudolf Kerschbamer, and Matthias Sutter, 2013, What Drives Taxi Drivers? A Field Experiment on Fraud in a Market for Credence Goods, *The Review* of Economic Studies 80, 876–891.
- Beasley, Mark S., 1996, An Empirical Analysis of the Relation between the Board of Director Composition and Financial Statement Fraud, *The Accounting Review* 71, 443–465.
- Benmelech, Efraim, and Carola Frydman, 2014, Military CEOs, *Journal of Financial Economics*.
- Bergstresser, Daniel, and Thomas Philippon, 2006, CEO incentives and earnings management, *Journal of Financial Economics* 80, 511–529.
- Bhojraj, Sanjeev, Paul Hribar, Marc Picconi, and John McInnis, 2009, Making sense of cents: An examination of firms that marginally miss or beat analyst forecasts, *The Journal of Finance* 64, 2361–2388.

Branigan, Tania, 2009, China to execute two over poisoned baby milk scandal, The Guardian.

- Brewer, Marilynn B., and Roderick M. Kramer, 1986, Choice behavior in social dilemmas: Effects of social identity, group size, and decision framing., *Journal of Personality and Social Psychology* 50.
- Burgstahler, David, and Ilia Dichev, 1997, Earnings management to avoid earnings decreases and losses, *Journal of Accounting and Economics* 24, 99–126.

- Burns, Natasha, and Simi Kedia, 2006, The impact of performance-based compensation on misreporting, *Journal of Financial Economics* 79, 35–67.
- Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler, 1997, Labor Supply of New York City Cabdrivers: One Day at a Time, *Quarterly Journal of Economics* 112, 407–441.
- Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller, 2011, Robust Inference With Multiway Clustering, *Journal of Business & Economic Statistics* 29, 238–249.
- Cameron, Jessica, and Dale Miller, 2009, Ethical standards in gain versus loss frames, in David De Cremer ed.: *Psychological Perspectives on Ethical Behavior and Decision Making* (Information Age Publishing, Charlotte, NC).
- Cameron, Jessica, Dale Miller, and Benoît Monin, 2010, Deservingness and unethical behavior in loss and gain frames. Working Paper, University of California, Berkeley and Stanford University.
- Chang, Otto H., and Joseph J. Schultz, 1990, The Income Tax Withholding Phenomenon: Evidence From TCMP Data, *Journal of the American Taxation Association* 12, 88.
- Christian, Charles, 1994, Voluntary Compliance with the Individual Income Tax: Results from the 1988 TCMP Study, *The IRS research bulletin*. Publication 1500 (U.S. Dept. of the Treasury, Internal Revenue Service, Washington, DC).
- Cohen-Cole, Ethan, 2011, Credit Card Redlining, *Review of Economics and Statistics* 93, 700–713.
- Dechow, Patricia M., Richard G. Sloan, and Amy P. Sweeney, 1996, Causes and Consequences of Earnings Manipulation: An Analysis of Firms Subject to Enforcement Actions by the SEC, *Contemporary Accounting Research* 13, 1–36.
- Dimmock, Stephen G., and William C. Gerken, 2012, Predicting fraud by investment managers, *Journal of Financial Economics* 105, 153–173.
- Efendi, Jap, Anup Srivastava, and Edward P. Swanson, 2007, Why do corporate managers misstate financial statements? The role of option compensation and other factors, *Journal of Financial Economics* 85, 667–708.
- Erickson, Merle, Michelle Hanlon, and Edward L. Maydew, 2006, Is There a Link between Executive Equity Incentives and Accounting Fraud?, *Journal of Accounting Research* 44, 113–143.
- Farber, David B., 2005, Restoring Trust after Fraud: Does Corporate Governance Matter?, *Accounting Review* 80, 539–561.
- Farber, Henry S., 2008, Reference-Dependent Preferences and Labor Supply: The Case of New York City Taxi Drivers, *The American Economic Review* 98, 1069–1082.

- Farber, Henry S., 2014, Why You Can't Find a Taxi in the Rain and Other Labor Supply Lessons from Cab Drivers. Working Paper, National Bureau of Economic Research.
- Gino, Francesca, and Lamar Pierce, 2010, Robin Hood Under the Hood: Wealth-Based Discrimination in Illicit Customer Help, *Organization Science* 21, 1176–1194.
- Grundfest, Joseph A., and Nadya Malenko, 2009, Quadrophobia: strategic rounding of EPS data, Rock Center for Corporate Governance at Stanford University Working Paper.
- Haddon, Heather, 2012, Another Side of Sandy Toll Seen in Taxis, Wall Street Journal.
- Haggag, Kareem, Brian McManus, and Giovanni Paci, 2014, Learning by Driving: Productivity Improvements by New York City Taxi Drivers. SSRN Scholarly Paper, Social Science Research Network, Rochester, NY.
- Haggag, Kareem, and Giovanni Paci, 2014, Default Tips, *American Economic Journal: Applied Economics* 6, 1–19.
- Hellevik, Ottar, 2007, Linear versus logistic regression when the dependent variable is a dichotomy, *Quality & Quantity* 43, 59–74.
- Jacob, Brian A., and Steven D. Levitt, 2003, Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating, *The Quarterly Journal of Economics* 118, 843–877.
- Jia, Yuping, Laurence Van Lent, and Yachang Zeng, 2014, Masculinity, Testosterone, and Financial Misreporting, *Journal of Accounting Research* 52, 1195–1246.
- John, Leslie K., George Loewenstein, and Scott I. Rick, 2014, Cheating more for less: Upward social comparisons motivate the poorly compensated to cheat, *Organizational Behavior and Human Decision Processes* 123. Behavioral Ethics, 101–109.
- Kahneman, Daniel, and Amos Tversky, 1979, Prospect Theory: An Analysis of Decision under Risk, *Econometrica* 47, 263–91.
- King, David A., and Jonathan R. Peters, 2012, Slow Down, You Move Too Fast: The Use of Tolls by Taxicabs in New York City, *Proceedings of the 91st Annual Meeting of the Transportation Research Board*.
- King, Gary, and Langche Zeng, 2001, Logistic Regression in Rare Events Data, *Political Analysis* 9, 137–163.
- Koch, Irmela, and Georg Wernicke, 2013, A special kind of CEO? Ex-military CEOs, financial fraud and stock option backdating, *Academy of Management Proceedings* 2013.
- Kranacher, Mary-Jo, Richard Riley, and Joseph T. Wells, 2010, *Forensic Accounting and Fraud Examination* (John Wiley & Sons).

- Lie, Erik, 2005, On the Timing of CEO Stock Option Awards, *Management Science* 51, 802–812.
- Malm, Sara, 2014, River of fire: Chinese waterway becomes so polluted it bursts into flame after lit cigarette is thrown into it, *The Daily Mail*.
- Morgenstern, Richard D., William A. Pizer, and Jhih-Shyang Shih, 2001, The cost of environmental protection, *Review of Economics and Statistics* 83, 732–738.
- Namako, Tom, 2010, Cabbies "Pass" up fast lanes to rip off riders, New York Post.
- Nelson, Mark W., John A. Elliott, and Robin L. Tarpley, 2002, Evidence from Auditors about Managers' and Auditors' Earnings Management Decisions, *The Accounting Review* 77, 175–202.
- NYC TLC, 2010, TLC Rules and Regulations: Chapter 54, Taxi and Limousine Commission, New York City.
- Orviska, Marta, and John Hudson, 2003, Tax evasion, civic duty and the law abiding citizen, *European Journal of Political Economy* 19, 83–102.
- Petersen, Mitchell A., 2009, Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches, *Review of Financial Studies* 22, 435–480.
- Porcano, Thomas M., 1988, Correlates of tax evasion, *Journal of Economic Psychology* 9, 47–67.
- Rhine, Sherrie L. W., and William H. Greene, 2006, The Determinants of Being Unbanked for U.S. Immigrants, *Journal of Consumer Affairs* 40, 21–40.
- Rhine, Sherrie L. W., William H. Greene, and Maude Toussaint-Comeau, 2006, The Importance of Check-Cashing Businesses to the Unbanked: Racial/Ethnic Differences, *Review of Economics and Statistics* 88, 146–157.
- Rick, Scott, and George Loewenstein, 2008, Hypermotivation, *Journal of Marketing Research* 45, 645–648.
- Schepanski, Albert, and David Kelsey, 1990, Testing for framing effects in taxpayer compliance decisions, *Journal of the American Taxation Association* 12, 60–77.
- Slemrod, Joel, Marsha Blumenthal, and Charles Christian, 2001, Taxpayer response to an increased probability of audit: evidence from a controlled experiment in Minnesota, *Journal of Public Economics* 79, 455–483.
- Tomz, Micahael, Gary King, and Langche Zeng, 2003, ReLogit: Rare Events Logistic Regression, *Journal of Statistical Software* 08.

- VanderWeele, Tyler J., and Ilya Shpitser, 2011, A New Criterion for Confounder Selection, *Biometrics* 67, 1406–1413.
- White, Halbert, 1980, A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity, *Econometrica* 48, 817–838.
- Zahra, Shaker A., Richard L. Priem, and Abdul A. Rasheed, 2005, The Antecedents and Consequences of Top Management Fraud, *Journal of Management* 31, 803–828.

Zitzewitz, Eric, 2012, Forensic economics, Journal of Economic Literature, 731-769.

Figure 1

This map of the New York City metro area illustrates the location of the Newark Liberty Airport, Staten Island, and the Verrazano-Narrows Bridge. The five boroughs of New York City are darkly shaded. These are areas where NYC taxi cabs can pick up passengers. The lightly shaded area to the west (left) is New Jersey, and the lightly shaded area to the east (right) is Long Island (NY). NYC cabs are not allowed to pick up fares in either New Jersey (including the Newark Airport) or Long Island, but they can drop off passengers in both New Jersey and Long Island on fares that originate in the five boroughs.



### Panel A

This panel reports the summary statistics for NYC taxi trips that cross the Verrazano-Narrows Bridge and end on Staten Island.

Variable	n	Mean	Std. Dev.	1st Quartile	Median	3rd Quartile
Cash Lane at Toll (Dummy)	30,341	0.01	0.08	0	0	0
Prior use of Cash Lane at Toll	30,341	0.03	0.18	0	0	0
Fare (Dollars)	30,341	48.17	19.85	32	46	62.5
Tip (Dollars)	30,341	5.07	6.42	0	0	9.9
Time (Minutes)	30,341	37.34	16.36	26	35	45
Passengers (#)	30,341	1.85	1.48	1	1	2
Distance (Miles)	30,341	16.46	7.75	10.03	15.7	22.39
Cash Payment (Dummy)	30,341	0.42	0.49	0	0	1
Low Income (Dummy)	30,341	0.21	0.41	0	0	0
High Income (Dummy)	30,341	0.46	0.5	0	0	1
Traffic	30,341	7.83	3.7	5	7	10
Rush Hour (Dummy)	30,341	0.25	0.43	0	0	1
Winter (Dummy)	30,341	0.22	0.41	0	0	0

### Panel B

This panel reports OLS models estimating trip time and tip received as a function of control variables and a dummy variable for whether or not the cabbie used the cash lane at the Verrazano-Narrows bridge (connecting Brooklyn with Staten Island) for trips ending on Staten Island. Standard errors are clustered in two dimensions (Cameron, Gelbach, and Miller (2011)) at the driver and date level. \*, \*\*, and \*\*\* indicate statistical significance (2-tailed) at the p < 0.1, p < 0.05, and p < 0.01 levels, respectively. T-statistics are reported in brackets beneath coefficients.

Dependent Variable:	Time (Seconds)	Tip (Dollars)	Tip (Dollars)
Model:	1	2	3
Cash Lane at Toll	103.0*	1.0091**	1.1949**
	[1.755]	[2.4648]	[2.0213]
Time (Minutes)		0.0252***	0.0262***
		[9.5776]	[4.8916]
Fare (Dollars)		0.0587***	0.1377***
		[10.4129]	[11.8587]
Passengers (#)	1.257	-0.1141***	-0.1532***
	[0.407]	[-6.6327]	[-5.6575]
Distance (Miles)	78.78***	0.0480***	0.0572**
	[90.22]	[3.9072]	[2.4120]
Cash Payment (Dummy)	38.53***	-8.9568***	
	[3.859]	[-156.9485]	
Low Income (Dummy)	-27.36**	0.0413	0.1069
	[-2.110]	[0.6197]	[0.9850]
High Income (Dummy)	-47.05***	0.1383**	0.2006**
	[-4.239]	[2.3178]	[2.0518]
Traffic	20.60***	0.0054	0.0193
	[10.43]	[0.6883]	[1.6384]
Rush Hour (Dummy)	347.7***	-0.3994***	-0.3529***
	[23.64]	[-6.4301]	[-3.5531]
Winter (Dummy)	7.215	0.3529***	0.5606***
	[0.381]	[4.7806]	[5.0750]
Constant	532.7***	4.2383***	0.1189
	[18.60]	[32.7932]	[0.5868]
W/ 1.1 0.1 0			
Weekday fixed-effects	Yes	Yes	Yes
Observations	30,341	30,341	17,634
R-squared	0.419	0.526	0.2826

This panel reports the results of a logistic regression that predicts the occurrences of taxi drivers taking the cash toll lane on the Verrazano-Narrows Bridge (for trips that end in Staten Island). Taking the cash lane is fraudulent, and drivers are required to use the E-Z Pass lane. Standard errors are clustered at the driver level. \*, \*\*, and \*\*\* indicate statistical significance (2-tailed) at the p < 0.1, p < 0.05, and p < 0.01 levels, respectively. Z-statistics are reported in brackets beneath coefficients.

Dependent Variable:	Cash Lane at Toll (Dummy)
Prior Cash Lane at Toll	2.4718***
	[14.1404]
Time (Minutes)	-0.0131
	[-0.2276]
Fare (Dollars)	0.024
	[1.3130]
Passengers (#)	-0.0333
	[-0.9305]
Distance (Miles)	-0.0001
	[-0.0086]
Cash Payment (Dummy)	-0.2813**
	[-1.9907]
Low Income (Dummy)	0.6083***
	[3.0476]
High Income (Dummy)	0.1245
	[0.6490]
Traffic	-0.1316
	[-0.7987]
Rush Hour (Dummy)	-0.0093
	[-0.4985]
Winter (Dummy)	-0.2902
	[-1.5349]
Constant	-5.8623***
	[-17.4934]
Weekday fixed-effects	Yes
Observations	30,341
Pseudo R-squared	0.0761

### Panel A

This panel reports the summary statistics for NYC taxi trips that cross the Verrazano-Nar	rows
Bridge and end at the Newark Liberty Airport.	

Variable	n	Mean	Std. Dev.	1st Quartile	Median	3rd Quartile
Cash Lane at Toll (Dummy)	5,473	0.92	0.26	1	1	1
Prior use of Cash Lane at Toll	5,473	0.6	0.49	0	1	1
Fare (Dollars)	5,473	70.85	11.58	65	68.5	73.5
Tip (Dollars)	5,473	9.2	7.97	0	10	16.1
Time (Minutes)	5,473	37.60	12.79	29	34	43
Passengers (#)	5,473	1.65	1.2	1	1	2
Distance (Miles)	5,473	18.56	4.56	16.9	17.9	19.6
Cash Payment (Dummy)	5,473	0.3	0.46	0	0	1
Low Income (Dummy)	5,473	0.17	0.37	0	0	0
High Income (Dummy)	5,473	0.44	0.5	0	0	1
Traffic	5,473	8.04	3.77	5	7	10
Rush Hour (Dummy)	5,473	0.27	0.44	0	0	1
Winter (Dummy)	5,473	0.28	0.45	0	0	1

### Panel B

This panel reports the results of a logistic regression that predicts the occurrences of taxi drivers taking the cash toll lane on the Verrazano-Narrows Bridge (for trips that end in Staten Island or at the Newark Liberty Airport). Taking the cash lane is fraudulent, and drivers are required to use the E-Z Pass lane. Standard errors are clustered at the driver level. \*, \*\*, and \*\*\* indicate statistical significance (2-tailed) at the p < 0.1, p < 0.05, and p < 0.01 levels, respectively. Z-statistics are reported in brackets beneath coefficients.

Dependent Variable:	Cash Lane at Toll (Dummy)
<b>-</b>	
Destination: Newark Airport	7.6510***
	[24.5051]
Prior Cash Lane at Toll	3.0720***
	[27.0676]
Time (Minutes)	0.0027
	[0.4275]
Fare (Dollars)	0.0061
	[0.4530]
Passengers (#)	0.011
	[0.2672]
Distance (Miles)	-0.1297***
	[-4.8329]
Cash Payment (Dummy)	-0.2369***
	[-2.7408]
Low Income (Dummy)	0.1048
	[0.8242]
High Income (Dummy)	-0.1123
	[-0.9478]
Traffic	0.0026
	[0.2045]
Rush Hour (Dummy)	-0.1591
	[-1.6355]
Winter (Dummy)	0.0503
	[0.5253]
Constant	-3.7463***
	[-16.5387]
Weekday fixed-effects	Yes
Observations	35,814
Pseudo R-squared	0.859

### Panel A

This panel reports the summary statistics for NYC taxi trips that cross the Verrazano-Narrows Bridge and end in New Jersey at locations other than the Newark Liberty Airport.

Variable	n	Mean	Std. Dev.	1st Quartile	Median	3rd Quartile
Cash Lane at Toll (Dummy)	662	0.57	0.5	0	1	1
Prior use of Cash Lane at Toll	662	0.31	0.46	0	0	1
Fare (Dollars)	662	93.79	26.82	75	95	122.5
Tip (Dollars)	662	10.48	9.64	0	10	19.4
Time (Minutes)	662	46.08	22.64	28	41	62
Passengers (#)	662	1.45	0.78	1	1	2
Distance (Miles)	662	23.43	11.66	14.17	19.49	37.11
Cash Payment (Dummy)	662	0.2	0.4	0	0	0
Low Income (Dummy)	662	0.21	0.4	0	0	0
High Income (Dummy)	662	0.42	0.49	0	0	1
Traffic	662	8.03	3.61	5	8	10
Rush Hour (Dummy)	662	0.31	0.46	0	0	1
Winter (Dummy)	662	0.25	0.43	0	0	1

Panel B

This panel reports the results of a logistic regression that predicts the occurrences of taxi drivers taking the cash toll lane on the Verrazano-Narrows Bridge (for trips that end in Staten Island or New Jersey, including the Newark Liberty Airport). Taking the cash lane is fraudulent, and drivers are required to use the E-Z Pass lane. Standard errors are clustered at the driver level. \*, \*\*, and \*\*\* indicate statistical significance (2-tailed) at the p < 0.1, p < 0.05, and p < 0.01 levels, respectively. Z-statistics are reported in brackets beneath coefficients.

Dependent Variable:	Cash Lane at Toll (Dummy)
Destination: Newark Airport	7.9038***
	[31.7292]
Destination: NJ (non-Newark Airport)	6.5140***
	[17.1589]
Prior Cash Lane at Toll	2.9309***
	[20.7462]
Time (Minutes)	0.0042
	[1.0890]
Fare (Dollars)	-0.0088
	[-1.2289]
Passengers (#)	0.0235
	[0.6026]
Distance (Miles)	-0.0986***
	[-6.5385]
Cash Payment (Dummy)	-0.2623***
	[-3.2119]
Low Income (Dummy)	0.1331
	[1.1545]
High Income (Dummy)	-0.1416
	[-1.2676]
Traffic	0.0015
	[0.1272]
Rush Hour (Dummy)	-0.1205
	[-1.3495]
Winter (Dummy)	0.046
	[0.5028]
Constant	-3.5934***
	[-17.7653]
	X.
weekday fixed-effects	Yes
Ubservations	36,476
Pseudo R-squared	0.847

Table 5 reports the results of linear probability models that predict the occurrence of taxi drivers taking the cash toll lane on the Verrazano-Narrows Bridge (for trips that end in Staten Island or New Jersey, including the Newark Liberty Airport). Taking the cash lane is fraudulent, and drivers are required to use the E-Z Pass lane. Standard errors are clustered at the driver level. \*, \*\*, and \*\*\* indicate statistical significance (2-tailed) at the p < 0.1, p < 0.05, and p < 0.01 levels, respectively. T-statistics are reported in brackets beneath coefficients.

	Linear Prob. Model, DV: Cash Lane at Toll (Dummy)			
Dependent Variable:	Model 1	Model 2		
Destination: Newark Airport	0.8267***	0.7784***		
	[76.9493]	[39.4487]		
Destination: NJ (non-Newark Airport)	0.5183***	0.4736***		
	[20.1991]	[12.2494]		
Prior Cash Lane at Toll	0.1457***	-0.0525***		
	[17.9390]	[-5.5013]		
Time (Minutes)	-0.0002*	-0.0004		
	[-1.8820]	[-1.5472]		
Fare (Dollars)	0.0007**	0.0009		
	[2.4140]	[1.5455]		
Passengers (#)	0.0004	0.0063**		
	[0.6135]	[2.5712]		
Distance (Miles)	-0.0040***	-0.0046***		
	[-6.4038]	[-3.6317]		
Cash Payment (Dummy)	-0.0027*	0.0018		
	[-1.8833]	[0.6303]		
Low Income (Dummy)	0.0019	0.0072		
	[0.8809]	[1.3775]		
High Income (Dummy)	-0.0022	0.0011		
	[-1.1199]	[0.2603]		
Traffic	-0.0001	0.0005		
	[-0.6474]	[1.3196]		
Rush Hour (Dummy)	-0.0041**	-0.0018		
	[-2.2007]	[-0.5614]		
Winter (Dummy)	0.0003	0.0125***		
	[0.1790]	[3.5693]		
Constant	0.0463***	0.0554***		
	[11.2489]	[5.9604]		
Weekday fixed-effects	Yes	Yes		
Driver fixed-effects	No	Yes		
Observations	36,476	36,476		
R-squared	0.8564	0.947		

This table reports the results of logistic regressions that predict the likelihood of taxi drivers taking the cash toll lane on the Verrazano-Narrows Bridge (for trips that end in Staten Island or New Jersey, including the Newark Liberty Airport). Taking the cash lane is fraudulent, and drivers are required to use the E-Z Pass lane. To control for unbanked or very poor cab drivers affecting overall rates of E-Z Pass use, we include in these models a dummy variable that proxies for the past use of E-Z Pass by each driver (*Driver Used E-Z Pass for* >50% of tolls in prior month). Standard errors are clustered at the driver level. \*, \*\*, and \*\*\* indicate statistical significance (2-tailed) at the p < 0.1, p < 0.05, and p < 0.01 levels, respectively. Z-statistics are reported in brackets beneath coefficients.

Logistic Regression: dependent variable =1 if driver uses cash lane at toll booth					
	Model 1	Model 2	Model 3		
Driver Used E-ZPass for >50% of tolls in prior month	-3.8732***	-4.0524***	-4.0988***		
	[-12.0505]	[-10.5616]	[-11.0866]		
Destination: Newark Airport		7.6485***	7.9490***		
		[25.1003]	[31.0615]		
Destination: NJ (non-Newark Airport)			6.4412***		
			[16.4143]		
Prior Cash Lane at Toll	2.3212***	3.0205***	2.8905***		
	[12.2684]	[25.2885]	[19.7410]		
Time (Minutes)	-0.0004	0.0013	0.0037		
	[-0.0452]	[0.1970]	[0.8778]		
Fare (Dollars)	0.0213	0.0094	-0.0079		
	[1.3039]	[0.7666]	[-1.1343]		
Passengers (#)	0.0096	0.0095	0.0221		
	[0.1635]	[0.2207]	[0.5404]		
Distance (Miles)	-0.0233	-0.1318***	-0.0959***		
	[-0.7766]	[-5.6408]	[-6.7902]		
Cash Payment (Dummy)	-0.1987	-0.2279**	-0.2681***		
	[-1.3188]	[-2.4564]	[-3.0724]		
Low Income (Dummy)	0.6891***	0.1464	0.1655		
	[3.3220]	[1.0816]	[1.3465]		
High Income (Dummy)	0.2385	-0.072	-0.1077		
	[1.2327]	[-0.6023]	[-0.9429]		
Traffic	-0.0069	0.0079	0.0053		
	[-0.3701]	[0.6119]	[0.4487]		
Rush Hour (Dummy)	-0.1719	-0.1383	-0.1084		
	[-0.9837]	[-1.3678]	[-1.1653]		
Winter (Dummy)	-0.2257	0.1112	0.1135		
	[-1.1786]	[1.1307]	[1.2175]		
Constant	-2.2681***	0.062	0.2821		
	[-4.7931]	[0.1353]	[0.6586]		
Weekday fixed-effects	Yes	Yes	Yes		
Observations	27,728	32,942	33,548		
Pseudo R-squared	0.125	0.866	0.855		