Moral Hazard in Leasing Contracts: Evidence from the New York City Taxi Industry[1]

Henry Schneider[2]

Working Paper: November 2008

[1] I thank Dan Benjamin, Steve Berry, Joe Doyle, Kirabo Jackson, Fiona Scott Morton, Steve Nafziger, Michael Waldman, and various seminar participants for valuable comments, and Marci Meixler for able research assistance. I am grateful to Samara Epstein, Leah Flax, Eric Kim, and Andrew Salkin at the New York City Taxi and Limousine Commission, Bruce Schaller of Schaller Consulting, Regina Armstrong of Urbanomics, and numerous taxi drivers and fleet managers for data and discussions about the New York City taxi industry.

[2] Johnson Graduate School of Management, Cornell University, henry.schneider@cornell.edu

Abstract

In this study, I investigate the effects of moral hazard in leasing contracts by examining the driving outcomes of all long-term lessees and owner-operators of New York City taxis. I find that moral hazard explains a sizable fraction of lessees' accidents, driving violations, and vehicle inspection failures, and erodes a moderate fraction of industry income. To address the possibility of endogenous contract choice, I conduct an instrumental variables analysis on the cross-section of all drivers, and a panel-data analysis on a subset of drivers who switched from leasing to owning.

I Introduction

A wide range of goods, from commercial aircraft and medical equipment, to computers, vehicles, and livestock, have active leasing markets. In 2007, for example, $234 billion of the $846 billion of U.S. domestic business capital investment was financed through leasing (Global Insight, 2007), and the value of the rental stock of residential real estate was nearly $3 trillion. Given the benefits of leasing, this prevalence is unsurprising.[1] Yet basic contract theory raises the possibility of moral hazard whenever a lessee's actions affect a lessor's payoffs and these actions are hidden.

In this study, I provide evidence about the leasing moral hazard by examining the New York City taxi industry, which is split between taxis operated exclusively by lessees and taxis with owner-drivers. Lessees have significantly worse driving outcomes than owner-drivers: In 2005, long-term lessees experienced 62 per cent more accidents and 64 percent more driving violations per mile than owner-drivers, and operated taxis that failed vehicle emissions and safety inspections at a 67 percent higher rate.[2]

Moral hazard is an obvious candidate to explain these differences: Lessees pay less or none of many of the variable costs they generate, including for vehicle maintenance, repair, replacement, and insurance, and hence have incentives to choose inefficient levels of vehicle care and risk. First-best contracts are generally unavailable because a lessee's driving style cannot be monitored at reasonable cost. Contracting over driving outcomes instead of actions also faces obstacles since taxis are typically operated by multiple drivers, which prevents some driving outcomes (e.g., vehicle mechanical failures) from being matched to individual drivers, because of unattractive risk-sharing properties (lessees tend to have modest means and may be ill-equipped to handle income shocks), and because collection of damages by lessors from lessees is often difficult. Nevertheless, some safeguards against moral hazard do exist, including modest deposits in case of an accident (more details of the driving arrangements are provided in section II).

An obvious second explanation for lessees' inferior outcomes is that lessees are inher-

[1] A non-exhaustive list of leasing benefits discussed in previous work includes giving the primary seller control over secondary markets (Waldman (1997)), mitigating adverse selection in secondary markets (Hendel and Lizzeri (2002), Johnson and Waldman (2003)), reducing trading frictions (Gavazza (2007)), and providing tax advantages (Franks and Hodges (1987)).

[2] Taxis, however, are still safer than other automobile traffic in New York City, averaging 4.6 injury accidents per million miles versus 6.6 injury accidents for automobiles overall (Schaller, 2006b).

ently riskier drivers - for example, long-term lessees have eight fewer years of taxi-driving experience on average than owner-drivers - and hence selection may be important as well.

To identify the effects of moral hazard separately from selection, I analyze proprietary data provided by the New York City Taxi and Limousine Commission (TLC) on all 17,972 New York City taxi drivers operating as long-term lessees or owner-drivers during the spring of 2005. Since owner-driver outcomes are first-best in expectation, they provide a natural benchmark with which to compare lessee outcomes: After controlling for driver and vehicle characteristics, the expected lessee-owner outcome gap represents the effect of moral hazard.

Of course, controlling for driver and vehicle characteristics is not straightforward: As with most empirical work in contract theory, addressing the possibility of endogenous contract choice is a primary challenge.[3] I address this challenge in three ways. First, I estimate the difference in outcomes between lessees and owner-drivers conditioning on a rich set of observed driver characteristics. Second, I conduct an instrumental variables analysis to address the possibility of unobserved driving risk that is correlated with leasing choice, instrumenting for leasing choice with community norms for taxi-ownership. Third, I compare the before and after outcomes of the 1,130 drivers who switched from leasing to owning during the sample period, which ensures the absence of endogeneity from time-invariant unobserved factors. All of these approaches yield qualitatively similar results. Finally, I conduct two robustness checks. In the first, I demonstrate the absence of a primary endogeneity concern, the attrition of drivers from the industry based on their driving outcomes (and hence driving risk) before they switch from leasing to owning. In the second, I show that, as predicted, leasing strongly affects the violations rate for speeding, which is particularly susceptible to moral hazard, but not seatbelt or cell-phone infractions, which should be largely immune.

After controlling for vehicle usage and driver characteristics, I estimate that moral hazard explains 34 percent of lessees' violations, 18 percent of their accidents, and 30 percent of leased taxis' vehicle inspection failures. While data on the full set of costs associated with taxi operation are unavailable, I am able to derive an estimator of the social costs of moral hazard by imposing restrictions on the relationship between operating costs and revenues. While tentative in nature, I estimate an annual social cost of moral hazard

[3]See Chiappori and Salanie (2003) for a discussion.

of \$79 million, an amount that represents 18 percent of the \$450 million in cumulative net income of long-term lessors and lessees.

To the best of my knowledge, this is the first detailed empirical analysis of the leasing moral hazard, and I interpret the magnitudes of these effects as moderately large. Dunham (2003) provides related evidence, showing that company vehicles have low resale values, which is consistent with employees providing inferior care when companies bear these costs. However, data limitations prevent a range of alternative explanations, such as inferior company maintenance practices, from being ruled out. The leasing moral hazard also resembles the moral hazard in auto insurance, and the current study contributes to a growing empirical literature on this topic as well (e.g., Abbring et al. (2003), Israel (2007)).[4] These studies test for changes in driving outcomes in response to changes in insurance premiums, and find small or no moral hazard effects. The current study builds on this work in the following ways. First, it examines a setting with larger variation in incentives for good driving outcomes, which facilitates identification. In Israel (2007), for example, identification is based primarily on the effect of \$50 and \$150 changes in annual insurance premiums, whereas the differences in incentives faced by taxi lessees versus owner-drivers is much larger. Second, estimates of the effects of moral hazard on a more comprehensive set of outcomes are provided. Finally, the richness of the current data allows me to employ a range of estimation methods and robustness checks, which lends added confidence in the results. I interpret the current results as suggesting that moral hazard in auto insurance contracts may be more important that previously recognized.

The results also speak to an active empirical literature that takes the effects of incentives as given, and tests whether observed contracts and firm organizations conform to theoretical predictions (e.g., Laffont and Matoussi (1995) and Ackerberg and Botticini (1999) on sharecropping, Benjamin et al. (1998) on real estate leasing). The current results suggest that moral hazard is important enough to justify the presence of safeguards in many types of leasing contracts (e.g., deposits in real estate rentals, maintenance schedules in vehicle leasing), and some of the make-versus-buy patterns that arise when the efficient use of a firm's assets cannot be assured (e.g., Baker and Hubbard (2003, 2004) in the trucking industry, Rawley and Simcoe (2007) in the taxi and limousine industry).

[4]There exists a larger empirical literature on asymmetric information in insurance that does not attempt to identify moral hazard separately from adverse selection (e.g., Chiappori and Salanie (2000)), but there are important normative differences between the two that make separate identification desirable.

Finally, the results inform the debate about how to improve the levels of safety and vehicle emissions associated with taxi ridership. Since a moral hazard in the leasing contract explains much of lessees' inferior outcomes, policies targeting incentives are likely to be effective.

While this study focuses on the costs associated with leasing, it is important to note that the arrangement provides a range of valuable benefits as well, including making the driver the residual claimant to fare revenue (thereby incentivizing effort relative to a fixed-wage contract), lowering barriers into the industry for drivers who are otherwise unable to afford the large capital requirements of owning, and limiting the liability from bad outcomes for drivers who can least afford them. In fact, the widespread adoption of leasing itself indicates that the value of the arrangement often exceeds the substantial costs described in this paper, providing a testament to its benefits relative to available alternatives.

II Details of the driving arrangements

To operate a taxi in New York City, a driver must possess a taxi license (often called a medallion) that has either an individual designation, which requires the owner to drive at least 210 shifts of at least nine hours per year, or a corporate designation, which is free of owner driving requirements.[5,6] The number of licenses is essentially fixed and licenses are traded between private parties for hundreds of thousands of dollars. From this licensing structure has arisen three types of driving arrangements: Owner-driving, in which the driver owns the medallion and the vehicle; DOV-driving, in which the driver leases the medallion but owns the vehicle; and lessee-driving, in which the driver leases both the medallion and vehicle. Lessee-drivers operate under either long-term leases, which typically extend for months, or short-term leases, which are 12-hours long.

Responsibility for paying the costs of operation varies considerably across these arrangements. Lessors pays for vehicle inspection ($50 for inspection, free for first re-inspection, $35 for additional re-inspections), maintenance and repair ($4,000 to $10,000 per year), vehicle replacement (approximately $28,000), insurance ($7,000 to $13,000 per year), workers' compensation ($1,400 to $3,000 per year), licensing fees (approximately $1,000 per year), and fines for improper keep of the taxi (Urbanomics (2004)).[7]

[5] Much of the information in this section comes from personal discussions with industry participants.

[6] The owner-must-drive requirement was enstated in 1990.

[7] In 2005, the TLC required no-fault insurance coverage of $200,000 per driver or passenger and $50,000

Lessees pays for gas ($5,000 to $10,000 per year), tolls, parking tickets, fines for DMV and TLC violations, TLC fees (approximately $100 per year), and lease fees. The TLC sets a maximum lease fee, which industry participants indicate is respected in practice and binds for nearly every lessor and shift (TLC (2006)). Lessees often provide a modest deposit, which is sometimes forfeited to the lessor after an at-fault accident.[8] Owner-drivers pay both sets of costs above. DOV-drivers also pay both sets of costs, except for liability insurance and workers' compensation, which medallion-lessors are required to pay.

Two industry programs exist that provide some incentive for better driving outcomes. The Persistent Violators program (and similar Critical Drivers program) stipulates a 30-day license suspension for 6 or more DMV or TLC points, and license revocation for 10 or more points, per 15-month period (completion of an accident-prevention course deducts 2 points). The Safe Drivers program reduces the bodily injury and property damage portion of the insurance premium (paid by the lessor) by 7 percent in the first year and 15 per cent in the second year if all lessees have no at-fault accidents in the prior 39 months and 4 or less DMV points in the prior 24 months.

Under long-term leasing, the lessee keeps the taxi for the entire period, but must return it to the base every one or two weeks for vehicle maintenance. In case of a mechanical problem or accident during operation, a stand-by vehicle is typically provided immediately to limit lost driving time, a practice which limits a lessee's cost of bad outcomes. The TLC requires that taxis pass a vehicle emissions and safety inspection every four months. In 2005, failing vehicles were typically kept in operation until the scheduled reinspection, which limits the cost to lessors of poor vehicle condition.[9]

For owner-drivers who lease out their taxis for second-shifts, the lessor-lessee match often occurs through family members, neighbors, and friends, though anonymous matches are common. For lease-only taxis, a lessee typically chooses a lessor to which he has no personal connections, and provides his TLC license number, lease payment, and sometimes a deposit. While most lessors have immediate access to the lessee's DMV driving history through a TLC database, leases are typically granted to any driver with less than two

per pedestrian, and liability insurance coverage of $100,000, $300,000, and $50,000 for non-economic losses per person, per accident, and for property damage.

[8]In 2005, the TLC capped the deposit amount at $80 and $500 for short and long-term leases.

[9]Starting in 2006, the TLC required the previous inspection certification sticker be removed before inspection so that failing taxis in operation are at higher risk of detection, which has likely reduced this practice.

serious accidents during the prior four years, a term dictated by the insurance companies.

III The data

Data were provided to me by the TLC under a proprietary data-sharing agreement and contain information about all New York City yellow taxis and drivers through 2005. I supplement these records with Census data matched to drivers using their addresses.

For every driver, the TLC data record all accidents involving injuries or property damage exceeding $1000 to any vehicle involved, and all driving violations accrued while operating taxis. The most common violations during 2005 were "Disobeying Traffic Device" (32 percent of violations; 2 points), "Speeding" (14 percent; average of 4.9 points depending on speed), "Improper Turn" (11 percent; 2 points), "Passed Red Light" (8 percent; 3 points), "Failed to Stop at Sign" (5 percent; 3 points), and "Failed to Yield to Pedestrian" (4 percent; 3 points). Cell-phone and seatbelt violations are also common, but I exclude them from the total violations frequency because they do not reflect driving aggressiveness and no points are assigned. The data also describe a range of driver characteristics, from age and New York City taxi-driving experience, to country-of-origin and English-language and driving-test results.[10]

For every taxi, the data record whether the vehicle passed the mandated triannual emissions and safety inspection, the odometer reading at time of inspection, and the number of fare-generating trips during the prior four months. The data also identify the taxi that each owner-driver and long-term lessee is registered to operate on April 2, 2005, which I use to merge the driver and taxi-level data. Since I only know the driver-taxi match at this one point in time, I restrict the majority of the analysis to the six months surrounding this date, and assume the match is stable over this period. Since short-term lessees typically drive different vehicles every day or week, they cannot be matched to specific taxis, and are excluded from the analysis. In the conclusion, I argue that moral hazard is likely larger for these drivers.

While the data do not identify drivers' vehicle usage rates explicitly, the odometer readings allow me to calculate the current period accumulated mileage for all drivers of

[10]Taxi-driving experience is calculated as the number of months registered as an active New York City taxi-driver, which I derive from the driver's hack license number using an algorithm provided to me by Bruce Schaller.

the taxi jointly.[11] The data are also limited in that they identify only medallion owners and not vehicle owners. Virtually all owner-drivers own both the medallion and vehicle. However, some drivers that I classify as lessees may own their vehicle despite leasing their medallion (DOV-operators), and these drivers are clearly less susceptible to moral hazard. While TLC officials and other industry participants have informed me that no industry-wide statistics on the number of DOV-operators exist, I believe the fraction of such drivers is modest: Even if each of the 4,946 taxis operated by long-term lessees had a driver who owned the vehicle (a vast overestimate), only 4,946 of the 12,430 long-term lessees (40 percent) would own their vehicle. In the conclusion, I discuss how the classification of DOV-operators as lessees will only generate conservative estimates of the leasing moral hazard.

Summary statistics

Table 1 provides summary statistics about the drivers examined in this study. There were 17,972 drivers registered as operators of specific taxicabs in the spring of 2005. Of these, 1,884 drivers are excluded from the analysis because key data fields are missing or contain original data-entry error. An inspection of these exclusions indicates that they are random in nature.

Of the included drivers, 12,430 were long-term lessees, while 3,658 were owner-drivers. As expected, owner-drivers tend to be older than lessees (51 versus 43 years old), and have more experience driving New York City taxis (18 versus 10 years). Nevertheless, considerable heterogeneity in age and experience exists for both lessees and owner-drivers, providing sufficient variation to identify their effects separately from the leasing effect. Figure 1 demonstrates this heterogeneity in histograms of experience for both lessees and owner-drivers.

Collectively, the 16,088 drivers lived in 6,707 Census blocks in the New York City area. Each driver lived in a Census block that had an average of 5.4 other active taxi drivers. As expected, owner-drivers lived in wealthier neighborhoods, as measured by vehicles per capita (0.33 versus 0.24), median income ($44,937 versus $38,038), median rent ($759 versus $693), and residents living in owner-occupied housing (46% versus 28%).

[11] For 15 percent of taxis, the current period mileage is unavailable, often because the vehicle was retired and replaced during the period. In such cases, previous period mileage is used instead. Estimates from a sample the excludes these taxis are nearly identical to the estimates from the full sample.

Over 90 percent of active taxi drivers were born outside of the United States. Table 2 lists driver statistics by country-of-origin, and shows that Pakistan, Bangladesh, India, and Haiti account for over half of all drivers. It is also apparent that lease rates vary widely across country-of-origin. For example, 98 percent of Senegalese and Moroccan drivers lease their taxi, whereas only 47 percent of Haitian drivers and 32 percent of Chinese drivers lease their taxis.[12]

In the spring of 2005, 12,779 taxis were licensed by the TLC for operation. Of these, 9,535 were operated by long-term lessees or owner-drivers, while 3,244 were operated by short-term lessees. Of the 9,535 long-term lessee and owner-driven taxis, 9,081 had sufficient data to include in my analysis, of which 2,008 were operated exclusively by owner-drivers, 2,040 had both owner-drivers and long-term lessees, and 4,994 were operated exclusively by long-term lessees. Among long-term lessee taxis, 39 percent were shared by two drivers, while 46 percent had three or more drivers.

Notable differences between owner and lessee-driven cabs are apparent from Table 3. Owner-only cabs were driven less than lessee-only cabs (118 versus 191 miles per day) and had lower odometer readings (118,328 versus 156,521 miles). These differences arise primarily because owner-only taxis are more likely to be driven two shifts per day (owner-only taxis had 1.1 drivers per taxi, while lessee-only cabs had 2.5 drivers per taxi). Drivers of lessee-only taxis had 62 percent more accidents and 64 percent more violations per mile than drivers of owner-only taxis (4.0 versus 2.5 accidents per million miles; 7.5 versus 12.3 violations per million miles). In addition, lessee-only taxis passed inspection on the first attempt 38 percent less often than owner-only taxis (40 versus 64 percent).

Figure 2 shows unconditional accident and violations rates by experience and driving arrangement, and shows that driving outcomes improve dramatically with experience for both lessees and owner-drivers. Owner-drivers, however, have better outcomes at all levels of experience. The figure also shows that violations are approximately three times more frequent than accidents at all levels of experience, foreshadowing that there will be more power to identify the violations model.

[12] At the request of the TLC, driving outcomes by country-of-origin are not reported.

Vehicle condition and accident frequency

Before presenting the empirical models of driving violations and accident frequencies, I first provide evidence that vehicle condition does not affect accident frequency, and hence can be excluded from the models (driving violations frequency is similarly unaffected).

Every taxi receives a rigorous vehicle inspection (52 percent of taxis fail inspection on the first attempt), which includes a safety check of alignment, suspension, headlights, and brakes, approximately every 120 days. Since any defects that increase accident frequency would accumulate during the period between inspections, accident frequency should increase in the time since the last inspection if vehicle condition is important. Figure 3 is a histogram of accident frequency by the number of days since the previous inspection, and no increase in accident frequency is visible.[13] I also estimate a linear fit of accident frequency versus number of days since last inspection and find no relationship. Limiting the sample to taxis with higher odometer readings, and to taxis that failed the prior inspection, both of which may be more susceptible to defects, again shows no relationship between accident frequency and number of days since last inspection (these latter results are not reported).

IV Estimating the effect of moral hazard on violations and accidents

While accident frequency provides a direct measure of driving risk, driving violations frequency also reflects this risk. In fact, Gebers (2003) finds that prior violations more accurately predict future accidents than prior accidents themselves (which is also a strong predictor of future accidents).

Empirical model

The empirical distributions over drivers' driving violations and accident frequencies are characterized by many zeros and right-skewness, and incidences accumulate through a Poisson arrival-rate-like process. Therefore, I model the random variables governing driving violations and accident frequencies as Poisson processes. I provide test results later in this

[13] Taxis for which the vehicle was retired during the period (almost always because of the TLC-mandated retirement age) are excluded.

section in support of the Poisson specification.

The density of the number of violations or accidents experienced by driver j on taxi i, y_{ij}, is,

$$f(y_{ij}|\cdot) = \frac{e^{-\mu_{ij}}\mu_{ij}^{y_{ij}}}{y_{ij}!} \tag{1}$$

where the conditional mean, $\mu_{ij} = E[y_{ij}|\cdot]$, is assumed to have the following exponential form,

$$\mu_{ij} = m_{ij}\exp\{\alpha_0 + \alpha_1 L_{ij} + X_{ij}\alpha_2\} \tag{2}$$

The exposure variable, m_{ij}, is the driver's mileage during the first six-months of 2005, $L_{ij} = 1$ if the driver leases and $L_{ij} = 0$ if he owns, and X_{ij} is a row vector of driver characteristics determining violations frequency.

As mentioned in section III, mileage data are only available at the taxi-level and not the driver-level, which precludes direct estimation of equation 2. However, under reasonable conditions, which are described in the Appendix, the structural model in equation 2 is equivalent to the following reduced-form taxi-level model that requires only taxi-level mileage, m_i,

$$\mu_i = m_i\exp\{\alpha_0 + \alpha_1\overline{L}_i + \overline{X}_i\alpha_2\} \tag{3}$$

where μ_i is the conditional mean of taxi-level violations or accident frequency, and the overlines indicate mean values across the n_i drivers of taxi i.[14] The model in equation 3 is appealing because its coefficients represent both the effects of the mean characteristics of a taxi's drivers, and a close approximation of the structural parameters of equation 2.

Driving violations results

I estimate the model of driving violations in equation 3 using the Poisson maximum likelihood estimator. The dependent variable is number of driving violations incurred by all drivers of the taxi during the first six months of 2005, and estimates are reported in Table 4 as marginal effects calculated at the mean values of the driver characteristics, and the mean six-month taxi-level mileage (32,045 miles).

[14] For example, $\overline{L}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} L_{ij}$ and $\mu_i = E[\sum_{j=1}^{n_i} y_{ij}|\cdot]$. The outcomes of the 13 percent of drivers operating multiple taxis are distributed across the $K_j > 1$ taxis proportionally. I.e.,

$$y_{ij} = \frac{\left(\frac{m_i}{n_i}\right)y_j}{\sum_{k=1}^{K_j}\left(\frac{m_k}{n_k}\right)}$$

Column (1) contains estimates from a benchmark specification that excludes driver characteristics. As such, the lease estimate represents both the moral hazard and selection effects. Lessee-taxis incur 0.157 ($p<.001$) more violations per six-month period than owner-taxis. The specification in column (2) controls for experience and Census block-group income, and the leasing effect falls to 0.087 ($p<.001$). Driver age and fraction of Census block-group residents that are high-school graduates are added in the specification in column (3), but the leasing estimate is essentially unchanged. The last row of Table 4 contains the mean number of violations per lessee-only taxi during the first six months of 2005, 0.435. Dividing the estimate of the leasing effect by this amount indicates that approximately 21 percent of lessees' violations is attributable to moral hazard. Note that this estimate is per mile, and hence excludes any increase in violations from higher mileage by lessees due to moral hazard.

The estimate of the experience effect in column (3) indicates that an additional decade of experience reduces violations frequency by 0.081 ($p<.001$), which is approximately equal to the estimated leasing effect. The estimates also indicate that the violations rate is modestly lower for drivers from Census block-groups with higher income and education levels, and modestly higher for older drivers. The signs and magnitudes of these estimates all appear reasonable.

As discussed above, mileage is measured at the taxi-level and not the driver-level. Hence, estimates based on inter-taxi variation in leasing choice may be more reliable that estimates from intra-taxi variation. Column (4) contains estimates from a sample limited to taxis operated exclusively by lessees or exclusively by owners, which limits reliance on intra-taxi variation in leasing choice. The estimates are qualitatively unchanged.[15]

I estimated additional specifications that are not shown. Replacing the continuous forms of age and experience with indicators in two-year intervals had little effect on the leasing estimate. Driving and English-language test scores were available for drivers on 3,132 and 3,043 taxis respectively, and while both were weakly significant predictors of violations rate, their inclusion had little effect on the leasing estimate. Testing for the equality of the conditional mean and variance of the specification in column (3) reveals modest overdispersion ($\sigma^2 = 1.14\mu$). However, estimates from a negative binomial model, which allows for more flexibility in the conditional mean-variance relationship, returns

[15]For comparison, Column (7) reports OLS estimates of a linear version of the model in equation 3, where mileage enters additively and unconstrained. The leasing estimate is moderately larger.

virtually identical estimates and standard errors.

Accidents results

I now estimate the same specifications of the empirical model of driving outcomes as above (from equation 3), but with the dependent variable number of accidents over all drivers of the taxi during the first six months of 2005. Estimates are reported in Table 5 as marginal effects evaluated at the mean values of the regressors, and the mean six-month taxi-level mileage (32,045 miles).

The pattern of estimates from the accidents models mirrors those from the violations models, and so I limit discussion to the primary result of interest, the estimate of the leasing effect from the specification in column (3). It shows that leasing itself increases the number of accidents per taxi by 0.020 (p=.060) per six-month period. Dividing this amount by the mean accident frequency per lessee-only taxi of 0.129 (reported in the last row of Table 5) indicates that leasing itself accounts for approximately 15.5 percent of lessees' accidents.

Comparing the violations and accidents results reveals that the estimates from the accidents models are considerably less precise. A primary reason is that accidents occur far less frequently than violations, and hence the accident outcomes contain many more zeros: Only 12.3 percent of taxis had a driver with at least one accident versus 32.3 percent for violations.

I estimated additional specifications that are not shown. The leasing estimate was qualitatively unaffected by discretizing the experience and age variables, and by the inclusion of driving and English-language test results. Including beginning-of-period odometer value, a rough proxy for vehicle condition, corroborates the result in section III that vehicle condition has little effect on accident rate: The effect of odometer is approximately zero (p=.966). The conditional mean and variance of accident frequency are virtually identically ($\sigma^2 = 1.00\mu$) providing support for the Poisson model.

Instrumental variable analysis

Any component of driving risk, r, that is uncaptured by the observed driver characteristics in X and correlated with leasing choice, L, would appear naturally in the empirical model in equation 3 as,

$$\mu_i = m_i \exp\{\alpha_0 + \alpha_1 \overline{L}_i + \overline{X}_i \alpha_2\} \exp\{\overline{r}_i\} \tag{4}$$

If equation 4 is the true model but is estimated without accounting for $\overline{r}_i$, then the leasing effect would clearly suffer from omitted variable bias. One approach I employ to address this possibility is to instrument for leasing choice.

The nonlinear reduced-form relationship between $\overline{L_i}$ and $\overline{r_i}$ due to the discrete nature of leasing choice, and the non-additive separability of $\exp\{\overline{r}_i\}$, complicate the use of IV methods. As discussed in Mullahy (1997) and Wooldridge (2002), traditional nonlinear IV methods assume additive separability of the correlated unobserved term, and generally produce inconsistent estimates when this condition fails. Instead, I use the GMM-based IV estimator proposed in Mullahy (1997) that requires no assumptions about the reduced form of $\overline{L}_i$.

The instrument I employ for leasing choice is the fraction of other taxi drivers from the same country-of-origin who lease their taxis. Table 2 shows that lease rates vary widely across country-of-origin even after controlling for average driver age and experience. There are several reasons for this variation. First, until 2007, the TLC required legal permanent resident or citizenship status to become an owner, and resident status varies widely across country-of-origin.[16] It is also evident from Table 2 that drivers from predominantly Muslim countries have high lease rates. For example, 92 and 90 percent of drivers from Bangladesh and Pakistan lease compared to 69 percent of drivers from India. Islamic religious law forbids taking interest-bearing loans, which may promote a norm against taxi-ownership which typically requires such loans. Finally, there appears to simply exist idiosyncratic norms for leasing across countries. For example, drivers from Haiti lease at a far lower rate than drivers from neighboring Dominican Republic (47 versus 87 percent) despite having similar average ages and experiences.

Country-of-origin lease rates are a good predictor of the fraction of drivers on a taxi who lease. I estimate the linear projection,

$$\overline{L}_i = z_i \Pi + v_i \tag{5}$$

where z_i includes the mean of the observed driver characteristics by taxi, $\overline{X}_i$, the instru-

[16] For example, between 1991 and 2000, 297,154 Haitians and 251,519 Colombians received permanent resident or citizen status versus 26,974 Moroccans and 5,774 Senegalese (Department of Homeland Security, Yearbook of Immigration Statistics, 1996, 2006).

ment, which appears as the mean across drivers of the taxi, and the log of taxi mileage with its coefficient constrained to one to account for the exposure variable; and v_i is an independent error. The t-statistic for the instrument is 11.75, and a one standard deviation increase in the instrument predicts a 5.0 percent increase in the fraction of a taxi's drivers who lease.

For the instrument to be valid, however, it must also be exogenous in the structural equation. The IV estimator is clearly exogenous to any within-country variation in unobserved driving risk that is correlated with leasing choice (e.g., idiosyncratic preferences for aggressive driving). To violate exogeneity, then, there must exist cross-country variation in driving risk that is uncorrelated with cross-country variation in experience, age, income, and education that not only persists for many years after immigration (the majority of drivers have a decade or more of experience) but is also correlated with cross-country variation in lease rates. While I believe this risk is limited, I also estimate a specification that includes as a regressor the mean driving outcome across all other lessees from the same country-of-origin (excluding the lessee himself) to capture any cross-country variation in unobserved driving risk.

The IV Poisson results for violations are reported in columns (5) and (6) of Table 4.[17] The estimate of interest, the leasing effect from the specification that controls for mean violations frequency among other lessees from the same country-of-origin, is 0.149 (p=.063), indicating that leasing explains 34 percent of the 0.435 mean number of violations incurred by drivers of leased-taxis per six months. While this estimate lacks the precision of the non-IV estimates, it is somewhat larger and suggests that if anything unobserved driving risk causes the non-IV estimates to moderately understate the true effects of leasing incentives.[18]

The IV Poisson results for accidents are reported in columns (5) and (6) of Table 5. The estimate of the leasing effect from the specification that controls for the mean accident rate of other lessees from the same country-of-origin is 0.023 (p=0.630), which is close in magnitude to the non-IV estimates, though very imprecise. Nevertheless, taking

[17] I estimate the IV Poisson model in Stata using program code from Nichols (2007) modified to include an exposure variable. Standard errors are computed by nonparametric bootstrap with 400 replications.

[18] The IV leasing estimate in column (5) that excludes the average violations frequency of other lessees from the same country is even larger, though again imprecise. The 2SLS estimate is provided in column (8) for comparison, and is considerably larger than the IV Poisson estimate, though is based on a misspecified model in which mileage enters additively and separably.

this estimate at face value indicates that leasing itself explain 18 percent of the 0.129 mean number of accidents experienced by drivers of leased-taxis per six months.[19]

Analysis of switchers

Between 1999 and 2004, approximately 1,130 drivers purchased an individual medallion and switched from leasing to owning. A comparison of their outcomes from before and after switching is another approach to estimating the leasing effect, and avoids the possibility of endogeneity from time-invariant unobserved driver characteristics.

An analysis of the panel dataset of pre-2005 driver outcomes has several limitations. First, vehicle usage records are absent, and hence any systematic change in usage upon switching may bias the estimates of the leasing effect. Second, as discussed in section III, the data only record medallion and vehicle owners, and hence switchers are identified as drivers who switched from leasing to owning their medallion. While most switchers will have purchased their vehicle and medallion together, it is possible that switchers, more than medallion-lessees overall, owned their vehicle prior to switching (e.g., if they purchased a vehicle prior to and in anticipation of purchasing a medallion), which would lead to attenuated estimates of the leasing effect. Third, it is possible that switchers are systematically different than non-switchers, and hence the estimated leasing effect for switchers imperfectly represents the effect of switching for drivers who chose not to switch during the sample period. Finally, it is possible that some lessees adjust their driving styles in anticipation of switching or switch after a period of good outcomes (e.g., TLC rules require a non-suspended or revoked TLC drivers license to become an owner), which would also lead to attenuated estimates of the leasing effect. While I am unable to directly assess the importance of most of these factors, I believe that if anything their net impact will be to understate the leasing effect.

The large center panel of Table 7 contains the mean number of total violations, accidents, speeding violations, and seatbelt violations per driver for each of the six years before and after switching. Since the panel-data regression analysis below only employs drivers with variation in the outcome variable (driver fixed effects are included), only drivers with a positive occurrence of the outcome in at least one year are included in the tabulated

[19]The IV estimate of the leasing effect in column (5) is moderately larger still, while the 2SLS estimate in column (8) is substantially larger, though again is based on a misspecified model.

means.

Total violations, accidents, and speeding violations frequencies decrease noticeably over the period, with larger decreases near the switching period. Driving violations and speeding violations both fall by approximately 5 percentage points upon switching, though no decrease in accident frequency is apparent around the switching year. As a check, column (5) contains seatbelt violations frequencies, an outcome that should be insensitive to moral hazard, and indeed, no decrease in frequency is apparent between the pre and post-switching periods.[20]

More formal panel-data modeling of driving outcomes based on within-driver variation in outcomes over time is challenging: It is unclear whether the data contain sufficient variation to cleanly identify the leasing effect separately from time and outcome-year effects. For this reason, and the limitations of the data described above, I believe a visual inspection of the mean outcomes by year in the middle panel of Table 7 is somewhat more informative. Nevertheless, I estimate a panel-data model that includes the prior year's outcome to allow for the possibility that the previous year's outcome affects the current year's driving style and decision to switch, and outcome-year fixed effects to allow for the possibility that law enforcement or city driving conditions vary by year. I exclude outcomes for the switching year since drivers operate under both arrangements and driving style may represent neither.

Table 7 reports estimates from the following panel model of driving outcomes,

$$\mu_{jt} = \exp\{\eta_j + \delta_t + \gamma_1 L_{jt} + \gamma_2 x_{jt} + \gamma_3 y_{jt-1}\} \tag{6}$$

where j and t index driver and year, η_j are driver fixed effects, δ_t are year fixed effects, $L_{jt} = 1$ when driver j leases in year t and $L_{jt} = 0$ when he owns, x_{jt} is driving experience in year t, which is really a catch-all that will reflect age and other effects that change over time, and y_{jt-1} is the driving outcome from the previous year. Equation 6 is estimated using the standard fixed-effects Poisson estimator.

The leasing estimate for violations is 0.095 (p=.050), which is similar to the cross-sectional estimate from section IV. The leasing estimate for speeding violations is also positive, though imprecise. The leasing estimate for seatbelt violations is negative, though imprecise, which suggests that vehicle usage does not decrease upon switching, and hence that the positive estimates for violations are not due to lower usage upon switching. The

[20] Cell-phone use during operation was only outlawed in 2002, and hence insufficient data for panel analysis exists.

leasing estimate for accidents is small and negative, though too imprecise to draw meaningful conclusions.[21]

In summary, the pattern of switchers' outcomes are generally consistent with the cross-sectional results presented earlier, and provides reassurance of their validity. However, data limitations prevent more precise inferences about the leasing effects, especially for accidents, to be drawn from the panel of switchers' outcomes.

Two robustness checks

While the results from the non-IV, IV, and panel-data analyses are all broadly consistent, the TLC data permit additional robustness checks. First, I provide evidence for the absence of one obvious possible source of endogeneity, that drivers who experience bad outcomes drop out of the industry at a higher rate. Since drivers typically lease for many years before switching to owning, attrition based on outcomes raises the possibility that ownership is correlated with unobserved driving risk. I find no evidence of a relationship between outcomes and attrition: Of the 8,364 lessees with no driving violations during all of 2005, 59 percent were still active drivers at the end of October 2007, compared to 58 percent of the 2,918 lessees with at least one violation and 57 percent of the 547 drivers with at least 2 violations. Of the 9,425 lessees who had no accidents during 2005, 59 percent were still in the industry in October 2007, compared to 60 percent of the 1,077 lessees with one accident. Even among the 101 lessees with at least two accidents, 61 percent were still active taxi drivers in October 2007.

Second, as a general check of the empirical methodology, I estimate Poisson models of specific types of driving violations, comparing the leasing effect for speeding violations, which is highly susceptible to moral hazard, with the leasing effects for seatbelt and cell-phone violations, which should largely be immune.

The estimates from these models are provided in Table 8, and are reported as marginal effects evaluated at the means of the regressors. The estimate of the leasing effect on speeding violations indicates that leasing explains 0.025 ($p<.001$) of the 0.045 mean number of speeding violations per lessee-taxi during the first half of 2005. The leasing effect for seatbelt violations is less than half this amount and statistically indistinguishable from

[21] Estimates of the leasing effect are modestly larger for violations and modestly more negative for accidents when the lagged outcome variable is excluded from the model.

zero, and the leasing effect for cell-phone violations is approximately zero. This differential effect of leasing on speeding versus seatbelt and cell-phone violations provides additional support for the empirical methodology of this study.

This comparison also provides a check that any differences in the average time-of-day of operation between owner-drivers and lessees, which could be related to driving conditions, are not determining the results (e.g., if a higher fraction of taxis on the road between 1 and 5 A.M. are operated by lessees). If the estimated leasing effects actually represent a time-of-day effect, the estimates for driving violations should be similar to the estimates for seatbelt and cell-phone violations (which are not included as driving violations).

V Estimating the effect of moral hazard on vehicle condition

In addition to affecting driving outcomes, the leasing moral hazard may also lead to inferior vehicle conditions. Vehicle condition may be affected in two ways. First, since lessees generally do not pay for maintenance and repair, they may practice inefficiently low levels of vehicle care (e.g., accelerating and braking too aggressively). Second, since lessors do not bear the full costs of poor condition (e.g., a higher frequency of breakdown, which reduces lessee revenues), a double moral hazard may exist in which lessors provide inadequate maintenance.[22]

Empirical model

While a direct measure of vehicle condition is unavailable, the TLC data contain vehicle emissions and safety inspection outcomes for all taxis, which provide a good indication of vehicle condition. I assume inspection outcome is determined by the latent variable vehicle condition, c_i^*, such that,

$$c_i^* = \beta_0 + \beta_1 \overline{L}_i + \overline{X}_i \beta_2 + V_i \beta_3 + \epsilon_i \tag{7}$$

where $\overline{L}_i$ is the mean driving arrangement of the drivers on the taxi, $\overline{X}_i$ is a row vector of the mean characteristics of the drivers on the taxi, V_i is a row vector of vehicle characteristics, and ϵ_i is a random component with a standard normal distribution. Vehicle condition depends on driving arrangement to allow for the moral hazard that may cause lessees to

[22] Evidence presented in section III indicates that vehicle condition does not affect violations and accident frequencies, which implies that the double moral hazard is not relevant for those outcomes.

provide less vehicle care than owner-drivers, and a double moral hazard that may cause lessors to provide less maintenance than owner-drivers. Thus, the leasing estimate reflects both sources of moral hazard. Vehicle condition depends on driver characteristics to allow for the possibility of systematic differences in care across driver types.

This specification leads to the following index model of inspection outcomes,

$$prob[c_i = 1] = prob[c_i^* > 0] = \Phi[\beta_0 + \beta_1 \overline{L}_i + \overline{X}_i \beta_2 + V_i \beta_3] \tag{8}$$

where the indicator variable $c_i = 1$ if taxi i passes the inspection, and $c_i = 0$ if it fails, and $\Phi(\cdot)$ is the evaluation of the standard normal CDF, which implies the usual probit specification.

Inspections results

Table 6 contains estimates from the probit regression model in equation 8 reported as marginal effects evaluated at the means of the regressors. The dependent variable is an indicator for whether the taxi passed the spring-2005 inspection on the first attempt. Vehicle characteristics include odometer reading at time of inspection, miles accumulated since last inspection, and whether the vehicle is a Ford (91 percent are Fords, 6 percent are Toyotas, 2 percent are Hondas, and 1 percent are Isuzus). Driver characteristics include taxi-driving experience and Census block-group income. Other driver characteristics had little impact.[23]

The specification in column (1) includes only vehicle characteristics and the fraction of a taxi's drivers that lease. The estimate of the leasing effect is -0.199 (p<.001), which indicates that the leased-taxi pass rate is 19.9 percentage points lower than the owner-taxi pass rate after accounting for vehicle usage. The estimate of the effect of odometer indicates that the pass rate falls by 24 percentage points for each additional 100,000 miles (p<.001), though the estimate of the effect of miles accumulated since last inspection indicates that the pass rate is approximately 3 percentage points higher for each additional 10,000 miles (p<.001), perhaps indicating that vehicles in better condition are driven more. The estimate of the effect of vehicle type indicates that Fords are 13.9 percent less likely to pass inspection (p<.001) than the Japanese makes after controlling for vehicle usage.

[23]Inspection outcome data are unavailable for 1,961 taxis because of missing data, often because the vehicle was retired and replaced during the period and no inspection was conducted, or original data entry.

The specification in column (2) shows that driver experience and income have limited effects on inspection outcomes: An extra decade of experience decreases the pass rate by 3 percentage points (p=.004), while an extra $10,000 in Census block-group income increases the pass rate by less than 1 percentage point (p=.056). The estimate of the leasing effect is now modestly lower, indicating the leasing itself leads to a 17.1 percentage point lower pass rate (p<.001). Comparing this amount to the leased-taxi pass rate of 40 percent, reported in the last row of Table 6, indicates that leasing itself explains 30 percent of leased-taxis' inspection failures. The specification in column (3) is estimated from a sample of taxis operated exclusively by lessees or owners to limit reliance on intra-taxi variation, while the specification in column (4) contains estimates of a linear version of the model estimated with OLS. In both cases, the estimates are qualitatively unchanged.

VI Social cost of moral hazard

Having identified the effects of moral hazard on a range of important outcomes, I now examine the resulting social costs. This assessment is inherently tentative as an explicit calculation of costs requires substantially more data than is available. Nevertheless, by imposing restrictions on the relationship between driving revenues and costs, I derive an estimator of the social costs that are borne by lessees and lessors, and, to the extent they are reflected in taxi insurance premiums, by others affected by taxi accidents. Notable components of costs that are absent from this calculation include the inferior health of city residents from higher emissions of leased taxis (reflected in their higher inspection failure rates), and the effect on passenger service from more aggressive driving (which could be positive from reduced travel time or negative from a less comfortable ride).[24]

Theoretical framework and estimator

Drivers vary in their ability, θ, to minimize variables costs (e.g., higher-ability drivers are better at avoiding accidents). Each driver selects a level of driving aggressiveness, a, which is measured as expected revenue in dollars, such that expected revenue increases one-for-

[24] Also note that these estimates reflect the social costs that arise from the lessor's inability to perfectly align the lessee's incentives with her own, and obviously do not imply that leasing is a net negative compared to a counterfactual of abolishing leasing.

one with a.[25] Let an owner-driver's expected annual payoff from driving one 12-hour shift per day be denoted by,

$$\pi_o(a;\theta) = \begin{cases} a - \frac{1}{2}c_l\,(a-\theta)^2 - \frac{1}{2}c_o\,(a-\theta)^2 - p_o & \text{if } a \geq \theta \\ a - p_o & \text{if } a < \theta \end{cases}$$

which is revenue less a quadratic cost of aggressive driving that is borne by the driver regardless of ownership status (the c_l term; e.g., driving violations), a quadratic cost of aggressive driving that is borne through ownership (the c_o term; e.g., vehicle damage), and a price to own the taxi, p_o. Clearly $a > \theta$ at the optimal a. The owner-driver's first-order condition defines his optimal level of aggressiveness, a_o^*, which I will use as the social optimum,

$$\frac{\partial \pi_o}{\partial a} = 1 - (c_l + c_o)(a_o^* - \theta) = 0 \tag{9}$$

and hence, $a_o^* = \theta + \frac{1}{c_l + c_o}$. Let a lessee's expected payoff be denoted by,

$$\pi_l(a;\theta) = \begin{cases} a - \frac{1}{2}c_l(a-\theta)^2 - p_l & \text{if } a \geq \theta \\ a - p_l & \text{if } a < \theta \end{cases}$$

where p_l is lease price, and the first-order condition is,

$$\frac{\partial \pi_l}{\partial a} = 1 - c_l(a_l^* - \theta) = 0 \tag{10}$$

with solution, $a_l^* = \theta + \frac{1}{c_l} > a_o^*$. For an individual of type θ who is a lessee rather than an owner-driver, the annual social cost of moral hazard in dollars is then,

$$\pi_o(a_o^*;\theta) - \pi_o(a_l^*;\theta) = \frac{1}{2}\left(\frac{c_o}{c_l}\right)^2 \frac{1}{c_l + c_o} > 0 \tag{11}$$

Let $L = 1$ if the driver leases, and $L = 0$ if he owns. Then using the values of a_l^* and a_o^* calculated above, we can rewrite the optimal level of aggressiveness for a driver of type θ as,

$$a^* = \frac{1}{c_l + c_o} + L\left[\frac{1}{c_l} - \frac{1}{c_l + c_o}\right] + \theta \tag{12}$$

which translates directly into the linear regression equation,

$$a_{ij} = \delta_0 + \delta_1 L_{ij} + X_{ij}\delta_2 \tag{13}$$

[25] The payoff function can be easily modified to allow a to also reflect the driver's choice of the number of hours to operate during the 12-hour shift period without affecting the resulting social cost estimator.

where a_{ij} is measured as expected fare revenue for driver j on taxi i per 12-hour shift per day averaged over the year, and X_{ij} includes all determinants of type θ.

Notice that $\delta_1 = \frac{1}{c_l} - \frac{1}{c_l+c_o} = \left(\frac{c_o}{c_l}\right)\frac{1}{c_l+c_o}$. Hence, an estimator of the annual social cost of moral hazard per 12-hour lessee-shift in dollars, from equation 11, is,

$$\frac{\delta_1^2}{2\delta_0} \tag{14}$$

Social cost results

While driver-level data on fare revenue are unavailable, the TLC data record taxi-level number of trips taken with passengers over a four-month period in the spring of 2005. From these taxi-level trip data, I calculate fare revenue per 12-hour shift per day averaged over the year for taxi i as,

$$\overline{a}_i = \frac{\text{\# trips in 4 months for taxi } i \times \text{mean revenue per trip (\$10.34)}}{\text{\# 12-hour shifts per day (2)}} \times 3 \tag{15}$$

The mean New York City taxi trip fare of $10.34 is a 2005 estimate from Schaller (2006a) that is inclusive of surcharges and tips. To obtain δ_0 and δ_1 for equation 14, I then estimate the following modified version of equation 13 using OLS,

$$\overline{a}_i = \delta_0 + \delta_1\overline{L}_i + \overline{X}_i\delta_2 \tag{16}$$

where $\overline{L_i}$ and $\overline{X_i}$ are mean values across the n_i drivers of taxi i.

Table 9 contains estimates from the model in equation 16. The outcome of interest is the estimate of annual social cost per 12-hour lessee-shift derived from equation 14.[26] The estimate is $4,377 (s.e.=$330). As a point of comparison, lessors pay approximately $7,626 more per taxi for liability insurance and vehicle maintenance per year than owner-drivers Urbanomics (2004)), though note that this amount excludes several additional potential costs of moral hazard, such as a lower resale value upon mandated vehicle retirement and the double moral hazard that leads lessors to under-maintain their taxis, and benefits, most notably from the additional fare revenue from more aggressive driving. Aggregating the estimated cost per 12-hour lessee-shift across the 8,972 lease-only taxis (inclusive of short and long-term lessees) indicates an industry-wide annual social cost of $79 million.

[26] A possible concern is that owner-drivers choose to operate during more profitable times of the day, which would lead to an understated estimate of the leasing effect. Farber (2005) provides evidence that hourly fare revenue is roughly constant over the day, and in particular between day- and night-time shifts, likely because the number of drivers adjusts to daily fluctuations in the pattern of demand.

While the validity of the estimate of social costs clearly depends on a range of tentative functional form assumptions, \$79 million represents 18 percent of the approximately \$450 million cumulative net income that lessees and lessors received in 2003 Urbanomics (2004)), suggesting that moral hazard erodes a non-trivial fraction of industry income.

VII Concluding remarks

To the best of my knowledge, this is the first detailed empirical analysis of the effects of moral hazard in a leasing market. These effects appear to be important, and provide support for a basic prediction of contract theory. The results would also seem to justify the presence of safeguards against moral hazard, from security deposits to mandated maintenance schedules, that appear in many types of leasing contracts. However, these safeguards may only be partially effective, as taxi lessees often leave a modest deposit to cover vehicle damage, and are subject to license suspension for driving violations, and the effects still occur. The results also speak to an emerging empirical literature on moral hazard in auto insurance, which has found only limited effects. The current results suggest that financial incentives may in fact have an important effect on auto-insuree behavior.

There are reasons to believe the current estimates are conservative. First, data limitations require me to group together drivers who lease their medallion but own their vehicle, who are less susceptible to moral hazard, with drivers who lease their medallion and vehicle together. While I describe in section III why I believe only a small fraction of lessees are classified in this way, their presence will only lead to understated estimates of the leasing effect. Second, data limitations also prevent me from including short-term lessees. However, these lessees typically drive different taxis each day, and hence are less reliant on the condition of any single vehicle, which likely exacerbates the moral hazard for this group.

Appendix: Equivalence of driver and taxi-level models

I provide here the conditions under which the driver and taxi-level models of driving outcomes in equations 2 and 3 are equivalent. First let the outcomes of drivers operating the same taxi, conditional on the observables, be independent. Applying the fact that the sum of independent Poisson random variables is also a Poisson random variable with a mean equal to the sum of the individual means then allows equation 2 to be summed across all n_i drivers on taxi i,

$$\mu_i = \sum_{j=1}^{n_i} m_{ij} \exp\{\alpha_0 + \alpha_1 L_{ij} + X_{ij}\alpha_2\}$$

Then specify that all drivers operating the same taxi drive the same number of miles, such that $m_{ij} = \frac{m_i}{n_i}$,

$$\mu_i = \frac{m_i}{n_i} \sum_{j=1}^{n_i} \exp\{\alpha_0 + \alpha_1 L_{ij} + X_{ij}\alpha_2\}$$

It is not immediately obvious that this condition is harmless. However, section IV contains estimates based only on taxis operated exclusively by lessees or exclusively by owners, which limits relying on intra-taxi variation in leasing choice, and the results are essentially unchanged. Also note that the unit of observation is the taxi and not the driver, which also reduces reliance on intra-taxi variation.

Finally, applying a Taylor expansion shows that the exponential function is sufficiently linear over the ranges of values in X_{ij} to exchange the expectation and exponential functions with minimal error,

$$\frac{m_i}{n_i} \sum_{j=1}^{n_i} \exp\{\alpha_0 + \alpha_1 L_{ij} + X_{ij}\alpha_2\} \approx m_i \exp\{\alpha_0 + \alpha_1 \overline{L}_i + \overline{X}_i \alpha_2\}$$

The right-hand side of the equation is the taxi-level model in equation 3 containing the structural parameters from equation 2.

References

Abbring, Japp, Pierre-Andre Chiappori, James Heckman, and Jean Pinquet, "Adverse Selection and Moral Hazard in Insurance: Can Dynamic Data Help to Distinguish?," *Journal of the European Economic Association*, 2003, *1*, 512–512.

Ackerberg, Daniel and Maristella Botticini, "Endogenous Matching and the Empirical Determinants of Contract Form," *Journal of Political Economy*, 1999, *110* (3), 564–591.

Baker, George and Thomas Hubbard, "Make Versus Buy in Trucking: Asset Ownership, Job Design, and Information," *American Economic Review*, 2003, *93* (3), 551–572.

— **and** — , "Contractability and Asset Ownership: On-Board Computers and Governance in U.S. Trucking," *Quarterly Journal of Economics*, 2004, *119* (4), 1443–1479.

Benjamin, John, Kenneth Lusht, and James Shilling, "What Do Rental Contracts Reveal About Adverse Selection and Moral Hazard in Rental Housing Markets?," *Real Estate Economics*, 1998, *26* (2), 309–329.

Chiappori, Pierre-Andre and Bernard Salanie, "Testing for Asymmetric Information in Insurance Markets," *Journal of Political Economy*, 2000, *108* (1), 56–78.

_ **and** _ , "Testing Contract Theory: A Survey of Some Recent Work," in M. Dewatripont, L. Hansen, and S. Turnovsky, eds., *Advances in Economics and Econometrics*, Vol. 1, Cambridge University Press, 2003.

Dunham, Wayne, "Moral Hazard and the Market for Used Automobiles," *Review of Industrial Organization*, 2003, *23* (1), 65–83.

Farber, Henry, "Is Tomorrow Another Day? The Labor Supply of New York City Cabdrivers," *Journal of Political Economy*, 2005, *113* (1), 46–82.

Franks, Julian and Stewart Hodges, "Lease Valuation when Taxable Earnings are a Scarce Resource," *Journal of Finance*, 1987, *42*, 987–1005.

Gavazza, Alessandro, "Leasing and Secondary Markets: Theory and Evidence from Commercial Aircraft," 2007. Working paper.

Gebers, Michael, "An Inventory of California Driver Accident Risk Factors," *California Department of Motor Vehicles, Research and Development Branch*, 2003, *RSS-03-204.*

Global Insight, "U.S. Equipment Finance Market Study 2007-2008," 2007.

Hendel, Igal and Alessandro Lizzeri, "The Role of Leasing Under Adverse Selection," *Journal of Political Economy*, 2002, *110* (1), 113–143.

Israel, Mark, "Do We Drive More Safely When Accidents are More Expensive? Identifying Moral Hazard from Experience Rating Schemes," 2007. Working paper.

Johnson, Justin and Michael Waldman, "Leases, Lemons, and Buy-backs," *RAND Journal of Economics*, 2003, *34* (2), 247–265.

Laffont, Jean-Jacques and Mohamed Matoussi, "Moral Hazard, Financial Constraints and Sharecropping in El Oulja," *Review of Economic Studies*, 1995, *62* (3), 381–399.

Mullahy, John, "Instrumental-Variable Estimation of Count Data Models: Applications to Models of Cigarette Smoking Behavior," *The Review of Economics and Statistics*, 1997, *79* (4), 586–593.

Nichols, Austin, "ivpois: Stata Module for IV/GMM Poisson Regressions," 2007. http://ideas.repec.org/c/boc/bocode/s456890.html (March 20, 2008).

Rawley, Evan and Timothy Simcoe, "Horizontal Diversification and Vertical Contracting: Firm Scope and Asset Ownership in Taxi Fleets," 2007. Working paper.

Schaller, Bruce, "The New York City Taxicab Factbook," 2006. http://www.schallerconsult.com/taxi/taxifb.pdf (March 20, 1998).

_ , "Taxicab and Livery Crashes in New York City, 2004," 2006. http://www.schallerconsult.com/taxi/crash06.pdf (March 20, 2008).

TLC, "Leasing Review, Overview and Update," 2006. March 9 Commission Meeting Presentation.

Urbanomics, "Taxi & Limousine Commission Environmental Impact Statement Socioeconomic Conditions Report," 2004. Internal document prepared for the TLC.

Waldman, Michael, "Eliminating the Market for Secondhand Goods: An Alternate Explanation for Leasing," *Journal of Law and Economics*, 1997, *40*, 61–92.

Wooldridge, Jeffrey, *Econometric Analysis of Cross Section and Panel Data*, MIT Press, 2002.

Table 1: Driver-level summary statistics

	Long-term lessees	Owner-drivers
Number of drivers	12,430	3,658
Fraction female	0.005	0.007
Fraction white	0.28	0.38
New York City taxi-driving experience (years)	9.5	17.6
Age	42.8	51.2
Height (inches)	67.8	68.0
Body Mass Index	24.6	24.7
Driving Test Score (out of 100)	78.3	80.5
Fraction passing English-language test	0.92	0.89
Mean number of other taxi drivers in block	6.1	3.2
Fraction of other drivers in block who lease	0.66	0.42
Vehicles per capita in block-group	0.24	0.33
Median income in block-group ($)	38,038	44,937
Median rent in block-group ($)	693	759
Fraction residents in owner-occupied housing in block	0.28	0.46
Fraction immigrated in last five years in block-group	0.12	0.09
Fraction high-school grads in block-group	0.12	0.12

Note: The table contains summary statistics about the taxi drivers analyzed in this study. Neighborhood characteristics are derived from Census data on the residents of the drivers' block and block-group, derived from Census data matched to drivers' home addresses. Lessees indicate long-term lessees.

Table 2: Driver-level country-of-origin statistics

Country	N	Lease	Exp.	Age	Country	N	Lease	Exp.	Age
Afghanistan	95	0.68	14	44	Korea	87	0.40	16	57
Algeria	77	1.00	5	39	Lebanon	25	0.56	12	45
Argentina	20	0.65	18	60	Liberia	32	0.88	11	44
Bangladesh	2502	0.92	8	40	Mali	26	1.00	7	40
Bolivia	21	0.43	17	52	Mauritania	30	1.00	3	37
Brazil	38	0.63	12	49	Mexico	11	0.64	11	44
Bulgaria	34	0.68	12	44	Moldova	10	0.50	11	45
Burma	30	0.43	14	46	Morocco	274	0.98	6	38
Chile	11	0.45	20	55	Nepal	52	0.90	3	35
China	157	0.32	16	51	Nigeria	110	0.91	9	46
Columbia	231	0.59	21	56	Pakistan	2677	0.90	9	41
Cuba	10	0.50	29	63	Peru	74	0.61	19	54
Dom. Rep.	291	0.87	17	51	Poland	87	0.48	17	53
Ecuador	232	0.65	18	52	Puerto Rico	110	0.75	23	63
Egypt	672	0.87	10	43	Romania	207	0.62	13	50
El Salvador	28	0.46	16	50	Russia	242	0.57	13	50
Eritrea	17	0.88	19	50	Senegal	156	0.98	7	43
Ethiopia	68	0.96	13	43	Sierra Leone	21	1.00	4	35
Ghana	466	0.86	10	46	Spain	10	0.30	24	60
Greece	193	0.39	20	56	Sri Lanka	11	0.91	8	48
Guatemala	12	0.67	18	48	Sudan	74	0.97	6	41
Guinea	68	1.00	5	37	Syria	29	0.83	14	50
Guyana	48	0.48	13	46	Thailand	11	0.82	16	55
Haiti	1545	0.47	17	52	Tibet	10	1.00	2	30
Honduras	17	0.59	10	42	Togo	12	1.00	6	43
Hong Kong	59	0.36	17	48	Trinidad	48	0.50	19	55
Hungary	11	0.27	20	62	Tunisia	16	1.00	4	33
India	2141	0.69	10	42	Turkey	94	0.67	10	45
Indonesia	21	1.00	10	49	United States	446	0.65	17	52
Iran	26	0.54	15	53	Ukraine	65	0.46	12	52
Israel	46	0.35	24	56	Uruguay	10	0.70	11	51
Italy	17	0.35	22	61	Uzbekistan	43	0.65	8	44
Ivory Coast	95	0.96	7	39	Vietnam	56	0.45	16	46
Jamaica	81	0.70	18	55	Yemen	22	0.95	5	36
Jordan	16	0.81	9	43	Yugoslavia	23	0.39	13	51

Note: The table contains driver country-of-origin statistics for countries with at least 10 active long-term lessees or owner-drivers, including fraction leasing (Lease), mean experience (Exp), and mean age (Age).

Table 3: Taxi-level summary statistics

	Owner-taxis	Mixed-taxis	Lessee-taxis
N taxis	2,047	2,040	4,994
Mean N drivers per taxi	1.1	2.8	2.5
Mean experience (years)	17.8	12.4	8.9
Mean age	51.9	45.6	42.1
Miles per day (all drivers)	118	187	191
Accidents per 10^6 miles	2.48	3.64	4.02
Violations per 10^6 miles	7.49	11.48	12.32
Vehicle odometer	118,328	155,471	156,521
Vehicle inspection pass rate	63.8%	49.2%	39.6%

Note: The table contains spring 2005 summary statistics about the taxis analyzed in this study. Violations includes only infractions for which points are assigned (i.e., cell-phone and seatbelt violations are excluded).

Table 4: Models of violations frequency

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Poisson	Poisson	Poisson	Poisson	IV Pois.	IV Pois.	OLS	2SLS
Fraction lessees on taxi	0.157**	0.087**	0.090**	0.081**	0.178*	0.149	0.135**	0.272**
	(0.017)	(0.019)	(0.019)	(0.018)	(0.080)	(0.080)	(0.020)	(0.087)
Taxi-mean experience (x10)		-0.055**	-0.081**	-0.082**	-0.077**	-0.077**	-0.076**	-0.049*
		(0.011)	(0.013)	(0.014)	(0.027)	(0.025)	(0.013)	(0.020)
Taxi-mean Census income (x10^4)		-0.023**	-0.023**	-0.022**	-0.028**	-0.027**	-0.022**	-0.013*
		(0.005)	(0.005)	(0.005)	(0.008)	(0.008)	(0.004)	(0.006)
Taxi-mean age (x10)			0.029**	0.019	0.053**	0.048**	0.015	0.017
			(0.009)	(0.010)	(0.014)	(0.014)	(0.011)	(0.011)
Taxi-mean Census high school grads			-0.464**	-0.559**	-0.549**	-0.506**	-0.422**	-0.358*
			(0.162)	(0.173)	(0.232)	(0.247)	(0.162)	(0.168)
Taxi-mean violations by country						0.273*		0.162*
						(0.118)		(0.073)
Taxi mileage (x10^4)	–	–	–	–	–	–	0.047**	0.022
							(0.009)	(0.016)
N	9,081	9,081	9,081	7,041	9,081	9,081	9,081	9,081
Observed leased-taxi violations rate	0.435	0.435	0.435	0.435	0.435	0.435	0.435	0.435

Note: The dependent variable is number of violations for all drivers on the taxi during the first half of 2005. Estimates are reported as marginal effects. Column (4) contains estimates from a limited sample that excludes mixed lessee-owner taxis. Columns (7) and (8) are a linear version of the model estimated with OLS and 2SLS respectively. The effect of taxi mileage is constrained in the Poisson models, but additive and unrestricted in the linear models. 'Taxi-mean violations by country' is mean violations frequency of other drivers from the same country-of-origin averaged over drivers of the taxi. Estimates of the intercept are not reported. Standard errors are in parenthesis (robust standard errors in columns (7) and (8)). * and ** indicate significance at five and one per cent levels.

Table 5: Models of accidents frequency

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Poisson	Poisson	Poisson	Poisson	IV Pois.	IV Pois.	OLS	2SLS
Fraction lessees on taxi	0.047**	0.021*	0.020	0.018	0.028	0.023	0.034**	0.054
	(0.009)	(0.010)	(0.011)	(0.010)	(0.041)	(0.047)	(0.009)	(0.041)
Taxi-mean experience (x10)		-0.022**	-0.018*	-0.017*	-0.003	-0.002	-0.015**	-0.011
		(0.006)	(0.008)	(0.008)	(0.015)	(0.017)	(0.006)	(0.009)
Taxi-mean Census income (x$\$10^4$)		-0.008**	-0.008**	-0.008**	-0.011*	-0.012*	-0.007**	-0.006*
		(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.002)	(0.003)
Taxi-mean age (x10)			-0.004	-0.004	-0.006	-0.007	-0.010	0.010
			(0.005)	(0.006)	(0.007)	(0.007)	(0.005)	(0.005)
Taxi-mean Census high school grads			-0.000	0.030	-0.101	-0.082	0.012	0.028
			(0.086)	(0.091)	(0.126)	(0.124)	(0.082)	(0.084)
Taxi-mean accidents by country						0.137		0.099
						(0.129)		(0.055)
Taxi mileage (x10^4)	–	–	–	–	–	–	0.013**	0.017*
							(0.004)	(0.008)
N	9,081	9,081	9,081	7,041	9,081	9,081	9,081	9,081
Observed leased-taxi violations rate	0.129	0.129	0.129	0.129	0.129	0.129	0.129	0.129

Note: The dependent variable is number of accidents for all drivers on the taxi during the first half of 2005. Estimates are reported as marginal effects. Column (4) contains estimates from a limited sample that excludes mixed lessee-owner taxis. Columns (7) and (8) are a linear version of the model estimated with OLS and 2SLS respectively. The effect of taxi mileage is constrained in the Poisson models, but additive and unrestricted in the linear models. 'Taxi-mean accidents by country' is mean accident frequency of other drivers from the same country-of-origin averaged over drivers of the taxi. Estimates of the intercept are not reported. Standard errors are in parenthesis (robust standard errors in columns (7) and (8)). * and ** indicate significance at five and one percent levels.

Table 6: Models of vehicle inspection outcomes

	(1)	(2)	(3)	(4)
	Probit	Probit	Probit	OLS
Fraction lessees on taxi	-0.201**	-0.171**	-0.167**	-0.145**
	(0.018)	(0.02)	(0.022)	(0.017)
Odometer at inspection (10^5)	-0.239**	-0.241**	-0.257**	-0.208**
	(0.007)	(0.007)	(0.008)	(0.005)
Mileage since last inspection (10^4)	0.026**	0.030**	0.027**	0.025**
	(0.008)	(0.008)	(0.01)	(0.007)
Vehicle is Ford	-0.133**	-0.133**	-0.122**	-0.105**
	(0.023)	(0.023)	(0.027)	(0.017)
Taxi-mean experience (x10)		0.030**	0.031**	-0.024**
		(0.01)	(0.012)	(0.009)
Taxi-mean Census income (x$\$10^4$)		0.009	0.005	-0.077*
		(0.005)	(0.005)	(0.038)
N	7,574	7,574	5,783	7,574
Observed leased-taxi pass rate	0.40	0.40	0.40	0.40

Note: The dependent variable is equal to one if the taxi passed vehicle inspection on the first attempt in the spring of 2005, and zero if it failed. Estimates from the Probit models are reported as marginal effects. Column (3) contains estimates from a limited sample that excludes mixed lessee-owner taxis. Standard errors are reported in parenthesis (robust standard errors in column (4)). * and ** indicate significance at the 5 and 1 percent levels.

Table 7: Annual driving outcomes of switchers

	N Violations			N Accidents			N Speeding			N Seatbelt		
Year	N	Mean	SE	N	Mean	SE	N	Mean	SE	N	Mean	SE
-6	142	0.51	0.06	115	0.23	0.04	39	0.31	0.09	35	0.11	0.05
-5	302	0.54	0.04	226	0.22	0.03	108	0.23	0.04	84	0.19	0.05
-4	419	0.51	0.04	306	0.22	0.03	139	0.30	0.04	124	0.16	0.04
-3	594	0.46	0.03	447	0.22	0.02	191	0.20	0.03	167	0.11	0.02
-2	753	0.45	0.02	571	0.20	0.02	243	0.23	0.03	209	0.13	0.02
-1	866	0.42	0.02	664	0.17	0.01	283	0.20	0.02	246	0.20	0.03
0	905	0.37	0.02	686	0.16	0.01	288	0.16	0.02	261	0.30	0.03
1	892	0.37	0.02	660	0.17	0.01	278	0.14	0.02	255	0.17	0.03
2	760	0.36	0.02	567	0.15	0.01	234	0.09	0.02	223	0.15	0.02
3	557	0.32	0.03	432	0.10	0.01	160	0.04	0.02	156	0.17	0.03
4	371	0.33	0.03	299	0.12	0.02	118	0.09	0.03	96	0.26	0.05
5	231	0.30	0.03	195	0.10	0.02	74	0.12	0.04	63	0.25	0.06
6	95	0.38	0.06	81	0.12	0.04	35	0.11	0.05	31	0.23	0.09
Driver is lessee		0.095			-0.019			0.084			-0.211	
		(0.050)			(0.040)			(0.063)			(0.233)	
Experience (years x10)		-0.278**			-0.294**			-0.282			0.331	
		(0.105)			(0.089)			(0.152)			(0.512)	
Lag of dependent variable		-0.121**			-0.133**			-0.288**			-0.770**	
		(0.018)			(0.021)			(0.058)			(0.182)	
Driver, year fixed effects?		Yes			Yes			Yes			Yes	

Note: The top panel contains mean annual frequencies of total violations, accidents, speeding violations, and seatbelt violations per driver by year relative to switching year (Year). For example, -3 indicates three years prior to switching, and includes the 1998 outcomes of drivers who switched in 2001, the 1999 outcomes of drivers who switched in 2002, and so on. The adjacent column contains the standard error of the mean (SE). N is the number of observations from which the outcomes are calculated, which are the number of drivers with at least one positive outcome in any of the years. The bottom panel contains estimates from the conditional Poisson fixed-effects regression, with standard errors reported in parenthesis adjacent to the estimates. * and ** indicate significance at the five and one percent levels.

Table 8: Models of specific violations frequencies

	(1)	(2)	(3)
	Speeding	Cell Phone	Seatbelt
Fraction lessees on taxi	0.025**	0.001	0.012
	(0.006)	(0.007)	(0.007)
Taxi-mean experience (x10)	-0.003	0.000	0.011*
	(0.004)	(0.001)	(0.005)
Taxi-mean Census income (x10^4)	-0.002	-0.003	0.000
	(0.001)	(0.002)	(0.002)
Taxi-mean age (x10)	0.003	-0.002	-0.008*
	(0.003)	(0.004)	(0.004)
Taxi-mean Census high school grads	0.034	0.028	0.031
	(0.045)	(0.059)	(0.056)
Taxi mileage (x10^4)	–	–	–
N	9,081	9,081	9,081
Observed leased-taxi outcome rate	0.045	0.053	0.046

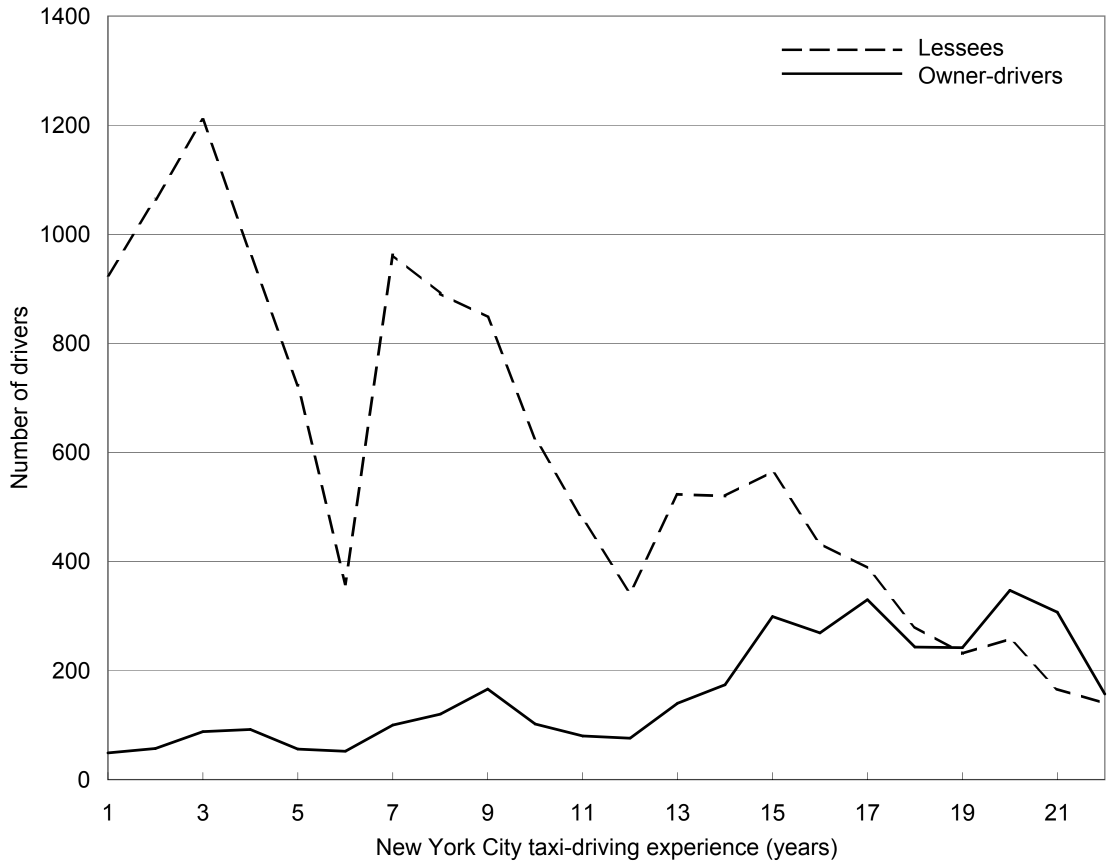
Note: The dependent variable is number of speeding, cell-phone, or seatbelt violations over all drivers on the taxi during the first six months of 2005. Estimates are from the Poisson model and reported as marginal effects. The effect of taxi mileage is constrained as described in the text. Estimates of the intercept are not reported. Standard errors are reported in parenthesis. * and ** indicate significance at five and one percent levels.

Table 9: Model of fare revenue

Fraction lessees on taxi (δ_1)	12,551**
	(354)
Taxi-mean experience (x10)	-96.2**
	(26.2)
Taxi-mean Census income (x$\$10^4$)	-803**
	(91.1)
Taxi-mean age (x10)	-179**
	(19.5)
Taxi-mean Census high school grads	4,456
	(3,322)
Intercept (δ_0)	35,992**
	(1,047)
Annual social cost per lessee-shift (\$) ($\frac{\delta_1^2}{2\delta_0}$)	4,377**
	(330)
N	6,687

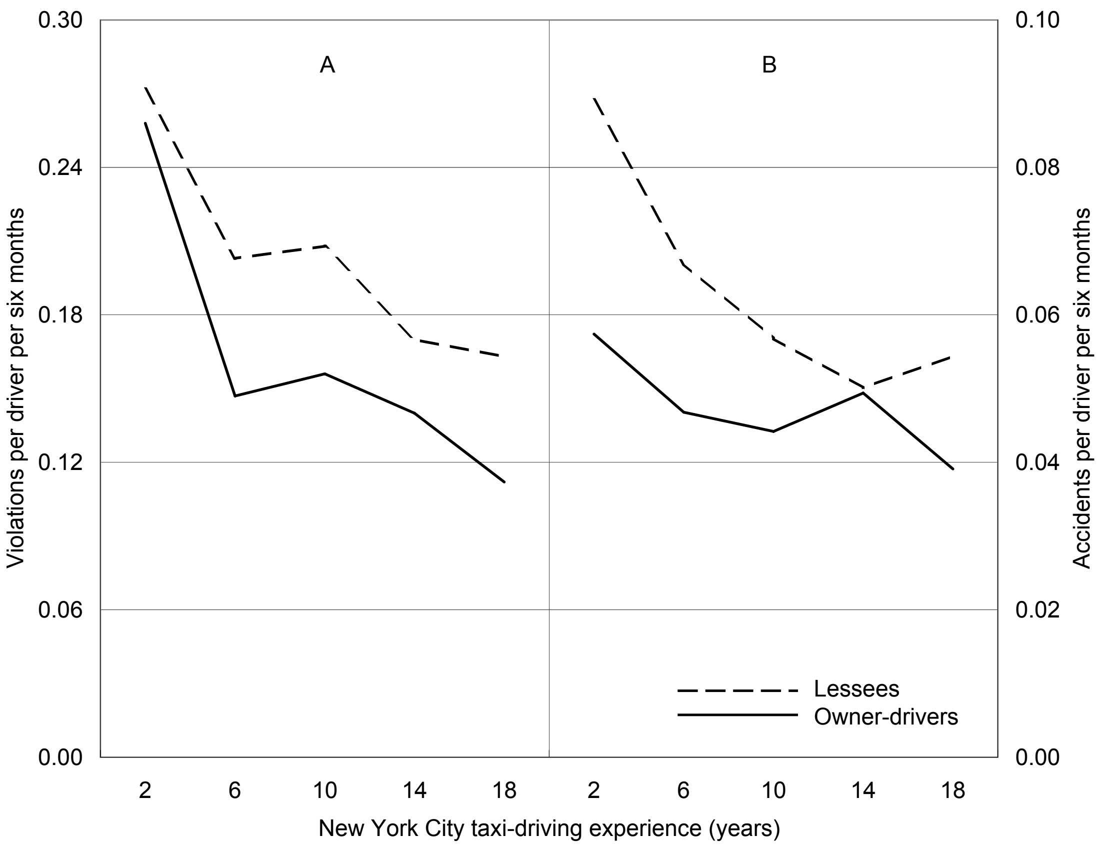
Note: The dependent variable is taxi-level fare revenue per 12-hour shift per day for all of 2005 (as described in section VI). The model is estimated with OLS. Robust standard errors are reported in parenthesis. * and ** indicate significance at five and one percent levels.

Figure 1: Histograms over experience



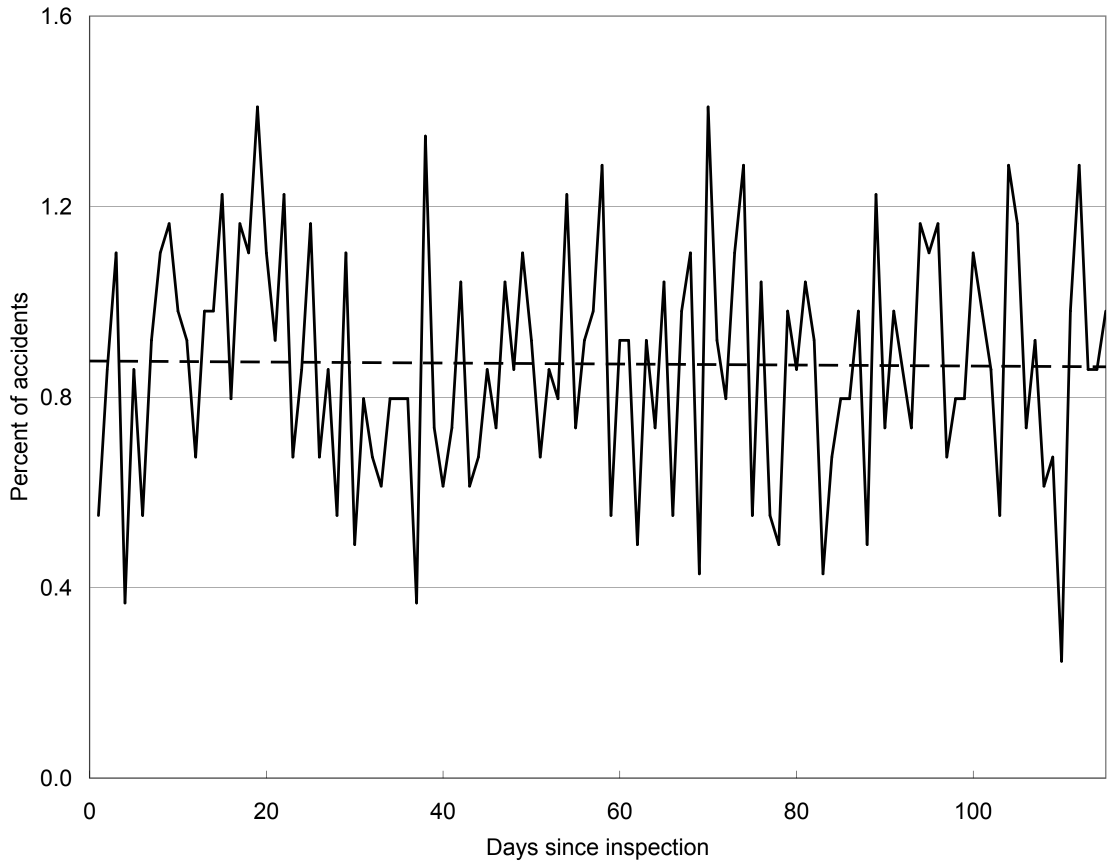
Note: The histogram shows the number of long-term lessees and owner-drivers by years of New York City taxi-driving experience in the spring of 2005.

Figure 2: Violations and accident frequencies by experience



Note: The figure shows the relationship between New York City taxi-driving experience and the number of violations (panel A) and accidents (panel B) per driver during the first six months of 2005. Frequencies are aggregated in four-year intervals around the labeled experience level to smooth the curves, and are unadjusted for taxi usage or other factors.

Figure 3: Accidents by days since inspection



Note: The curve shows the percent of total taxi accidents during the spring of 2005 by number of days since the last vehicle inspection was conducted. The dashed line is a linear fit of the curve.