2011

Do Social Connections Reduce Moral Hazard? Evidence from the New York City Taxi Industry

C. Kirabo Jackson, *Northwestern University*
Henry Schneider, *Cornell University*

American Economic Journal: Applied Economics 3 (*July 2011*)*: 244–267*
http://www.aeaweb.org/articles.php?doi=10.1257/app.3.3.244

Do Social Connections Reduce Moral Hazard? Evidence from the New York City Taxi Industry†

By C. Kirabo Jackson and Henry S. Schneider*

This study investigates the role of social networks in aligning the incentives of agents in settings with incomplete contracts. Specifically, the study examines the New York City taxi industry where taxis are often leased and lessee-drivers have worse driving outcomes than owner-drivers due to moral hazard. Using within-driver variation and instrumental variable strategies to remove selection, we find that drivers leasing from members of their country-of-birth community exhibit significantly reduced effects of moral hazard, representing an improvement of almost one-half of a standard deviation of the outcome measures. Screening is ruled out as an explanation, and other mechanisms are investigated. (*JEL* D82, D86, L92, Z13)

When an economic agent is not held accountable for the full cost of his or her actions, the situation may create a moral hazard; through doing what is personally optimal, the agent may behave in ways that are socially suboptimal (Bengt Holmstrom 1979). Moral hazard plays a central role in numerous contractual relationships, and researchers have found evidence of moral hazard in many contexts.[1] In this study, we investigate whether social ties can limit moral hazard in a setting where contracts are necessarily incomplete. Specifically, we examine the New York City (NYC) taxi-leasing market, comparing the outcomes of taxi drivers who lease from an owner from their same country of birth to the outcomes of drivers who lease from an owner from a different country of birth.

The taxi-leasing market is attractive for studying moral hazard because lessee-drivers pay little or none of many of the costs they generate, including those for vehicle maintenance and repair, and, hence, have incentives to choose inefficient levels of vehicle care. The challenge is that a driver's driving style is difficult to monitor, such that leasing contracts cannot fully specify the driver's level of vehicle care.[2] Schneider (2010) finds that a significant number of accidents and driving

* Jackson: Northwestern University, School of Education and Social Policy, 2120 Campus Drive, Room 204, Annenberg Hall, Evanston, IL 60208 (e-mail: kirabo-jackson@northwestern.edu); Schneider: 326 Sage Hall, Cornell University, Ithaca, NY 14853 (e-mail: henry.schneider@cornell.edu). We would like to thank Andrew Salkin at the New York City Taxi and Limousine Commission for helpful discussions and data, Ricard Gil, Jonathan Peterson, Lan Shi, Michael Waldman, two anonymous referees, and seminar participants at Cornell University, Depaul University, and University of Washington for valuable comments, and James Cowan and Jonathan Peterson for excellent research assistance.

† To comment on this article in the online discussion forum, or to view additional materials, visit the article page at http://www.aeaweb.org/articles.php?doi=10.1257/app.3.3.244.

[1] For example, worker effort is higher under piece rate versus hourly wage (Andrew D. Foster and Mark R. Rosenzweig 1994, Edward P. Lazear 2000), and make-versus-buy patterns in trucking may reflect these concerns (George P. Baker and Thomas N. Hubbard 2003, 2004).

[2] Drivers, however, keep all fare revenue, so an effort to generate fares for the owner is not relevant in this context.

violations for lessee-drivers is due to this moral hazard. This market is also attractive for studying social ties because country-of-birth networks have been found to be important in other settings, and most owners and drivers are immigrants to the United States.[3]

Our approach is to compare the driving outcomes of drivers who lease from owners from their same country of birth (i.e., in-network) to the outcomes of drivers who lease from owners from a different country of birth (i.e., out-of-network). The empirical difficulty in isolating the causal (effort) effect of in-network driving is that the outcomes of in-network drivers versus out-of-network drivers may either be due to effort effects (i.e., in-network drivers may exert less/more effort in their driving style to limit the costs borne by the owner because of greater potential social sanctions, social collateral, social preferences, or ease of monitoring) or due to differential selection into in-network leasing by driving ability (i.e., owners may be better positioned to identify and hire high ability drivers when they are in-network, or alternately, owners may be willing to hire lower ability drivers from in-network because they know that they will exert more effort). As such, our analysis focuses on isolating the causal (effort) effect of driving in-network.[4]

Because there is no experimental variation, our approach is not to provide a single estimate of the causal (effort) effect from a single research design, but to present a body of evidence based on several credible but imperfect research designs aimed at removing selection effects that rely on distinct sources of variation and different identifying assumptions. Our first strategy is to compare the *change* in outcomes of drivers who switch between in-network and out-of network leasing over time to the outcomes of drivers who always drive in-network or never drive in-network over the same period. This removes bias due to high ability drivers being more or less likely to lease in-network. Our second strategy exploits cross-sectional variation in the residential locations of drivers relative to owners *from the same country of birth*. We use the distance from a driver's residence to the nearest same-country owner's residence as an exogenous instrument for driving in-network, while also conditioning on the place-of-birth network. We present empirical tests to support the validity of both of these identification strategies. As supporting evidence, we then aggregate the data to the country level to compare driving outcomes for countries with high rates of in-network driving to outcomes for countries with low rates of in-network driving—removing selection bias *within* a country-of-birth network. While each of these three strategies may individually be subject to distinct sources of bias, it is unlikely that all three sources of variation are subject to bias. Thus, while none of these designs alone is dispositive, consistent results across approaches would indicate a real causal effect.

First, we show that in-network driving is much more common than would be expected by random chance and cannot be explained by residential clustering by

[3] For example, country-of-birth networks have been found to be important for immigrant assimilation (George J. Borjas 2000) and job seeking (Kaivan Munshi 2003; Per-Anders Edin, Peter Fredriksson, and Olof Åslund 2003). Mark Granovetter (1974), James D. Montgomery (1991), Roberto M. Fernandez and Nancy Weinberg (1997), and Michele Pellizzari (2010) studied the role of social networks in employment search.

[4] Note that we use the term *ability* broadly to indicate intrinsic driver characteristics that affect the driving outcomes of interest. These characteristics include motor skills and other factors, such as conscientiousness.

country of birth. This result is consistent with Ricard Gil and Wesley R. Hartmann (forthcoming), which found that markets may organize around common-origin communities to mitigate agency problems. Across all models, drivers have better outcomes when in-network than when out-of-network, indicating a positive effort effect. The effects are more negative in models that account for selection (relative to Ordinary Least Squares [OLS]), indicating *negative* selection into in-network driving, and suggesting that owners do not screen for high ability in-network drivers but, rather, hire lower ability in-network drivers because they exert more effort.

Second, we provide several robustness checks to present evidence against owner selection and to demonstrate that the largest effects are for summons types directly affecting owners' costs (as opposed to drivers' costs, which drivers would internalize). Third, we investigate the mechanisms through which social network effects operate. Previous work has proposed that contracts that exploit social ties via social sanctions and lower monitoring costs may reduce moral hazard (Joseph E. Stiglitz 1990). Dean Karlan et al. (2009) argue that network connections generate "social collateral" because the possibility of losing friendships secures informal transactions the same way that physical collateral can secure formal lending. Abhijit V. Banerjee, Timothy Besley, and Timothy W. Guinnane (1994); Besley and Stephen Coate (1995); and W. Bruce Wydick (1999) argue that it is the strength of group pressure on potential defaulters that matters, rather than social ties *per se*, while Alejandro Portes and Julia Sensenbrenner (1993) argue that altruism between members can align incentives. We find that network effects increase with owners' network strength but not drivers' network strength, which is consistent with the presence of social sanctions in the form of community-enforced punishments or social collateral. These would likely emanate from the owner's network.

Nearly all past research on social ties and moral hazard relates to group lending in developing countries, where markets are not well defined, credit is constrained, and property rights are not well enforced.[5] Moreover, these results are mixed.[6] As the first study to show directly the potential of social networks to reduce moral hazard in a developed nation and in a well-defined market, the current study adds breadth to the literature on social networks and moral hazard. That is, even in a setting in which credit is available to finance a taxi purchase, leasing is commonplace among members of the same country network, indicating that social networks significantly reduce the cost of using the market.[7]

[5] We are aware of only two empirical studies on social ties and moral hazard: Avner Greif (1993), who finds that community enforcement limited moral hazard in eleventh century Mediterranean trade; and Oriana Bandiera, Iwan Barankay, and Imran Rasul (2005, 2009), who find that social ties lead to lower productivity among college student fruit pickers.

[6] Researchers have argued that social ties help explain the success of group lending banks (Mark M. Pitt and Shahidur R. Khandker 1998, Banerjee et al. 2009). Michihiro Kandori (1992) and Glenn Ellison (1994) show theoretically how enforcement may operate, while Manohar Sharma and Manfred Zeller (1997), Karlan (2007), and Christian Ahlin and Robert M. Townsend (2007) show evidence for group lending, though Xavier Giné and Karlan (2006, 2009) find no benefits of group lending. Social interactions are also found to be economically relevant elsewhere (Marianne Bertrand, Erzo F. P. Luttmer, and Sendhil Mullainathan 2000; Edward L. Glaeser et al. 2000; Esther Duflo and Emmanual Saez 2003; Karlan 2005; Alexandre Mas and Enrico Moretti 2009; Jackson and Elias Bruegmann 2009).

[7] An active credit market exists in which specialized lenders market loans to taxi drivers to buy their own taxis.

The paper proceeds as follows. Section I describes the NYC taxi industry. Section II presents the data. Section III describes the empirical frameworks. Sections IV and V contain the results and robustness checks. Section VI explores mechanisms through which social network effects operate. Section VII concludes.

I. The New York City Taxi Industry

To operate a taxi in New York City, a driver must possess a taxi license, often called a medallion.[8] These licenses come in two forms: an "owner-must-drive" license, which requires the license owner to drive full-time (i.e., 210 shifts of at least 9 hours per year), and a corporate license, which is free of owner-driving requirements and typically is owned by corporations that lease the licenses (often with vehicles) to drivers. As such, drivers operate in one of three ways: owner-driving, in which the driver owns the license and vehicle; driver-owned-vehicle-driving, in which the driver leases the license but owns the vehicle; and lessee-driving, in which the driver leases both the license and vehicle. Lessee-drivers can operate under either long-term leases, which typically extend for 6 months at a time, or short-term leases, which are 12 hours long. Taxis are typically operated by two drivers for two shifts per day: a day shift and night shift. Thus, a taxi is typically operated by two lessee-drivers, two co-owner-drivers, or one owner-driver and one lessee-driver who leases from this owner-driver. For reasons explained below, we primarily examine this third type of driving arrangement, in which the lessee-driver leases the taxi (both the license and the vehicle) for one 12-hour shift per day from an owner-driver who operates the taxis for the second 12-hour shift per day.

The costs and revenues of operation are split between lessee-drivers and owner-drivers as follows. The owner is required by the NYC Taxi and Limousine Commission (TLC) to pay for no-fault and liability insurance and for workers' compensation for the lessee-driver, in addition to licensing fees, vehicle inspections, and fines for improper taxi use, some of which may be due to lessee actions. Owners also pay for all vehicle maintenance, repair, and replacement.[9] Lessees pay TLC fees and pay for their own gas usage, tolls, parking tickets, and some of their Division of Motor Vehicles (DMV) and TLC fines. Lessees also pay a flat lease fee to the owner. The lease fee is capped by the TLC at a maximum amount that owners can charge lessees per lease period, which industry participants indicate is respected in practice and nearly always binds. Lessee-drivers keep all fares and tips that they generate.

Long-term lessees often leave a modest deposit with the owner, which can be forfeited in the case of an at-fault accident. During the sample period, the deposit amount and lease fee are capped by the TLC at $500 and $650, respectively, which, as mentioned, industry participants indicate is respected. Note that these amounts are small compared to the cost of mechanical defects (approximately $7,000 per taxi/year on average), accident damage (up to $28,000 to replace the vehicle), and

[8] Much of the information in this section comes from personal discussions with industry participants, rules documents made available by the TLC, and Urbanomics (2004).

[9] Note that the vehicle owner pays all repair costs. Since we are primarily examining drivers who lease from owner-drivers, it is these owner-drivers who pay repair costs.

owners' insurance premiums (approximately $10,000 per taxi/year). Thus, while we cannot ensure that collateral and fee are uniform across drivers as we do not have these data, owners have limited leeway in varying contract terms relative to the operating costs.

Lessee-drivers generate three types of costs that are borne by owners. The first is repair costs from vehicle mechanical failures and maintenance. Given that NYC taxis are operated in congested urban conditions for 175 miles per day, on average, driver care can quickly show up in vehicle condition.[10] The second is from accidents. Approximately one-third of taxis are involved in a serious accident (defined as having over $1,000 in damage or involving injuries) per year. Driver aggressiveness, such as speeding or cutting across traffic to pick up a hailing passenger, increases the chance of accidents. The third is fines to owners resulting from driver behavior, ranging from $25 to $100 per instance, when drivers violate certain TLC regulations (discussed below). In other words, drivers pay only a fraction of the operating cost but keep all of the fare revenue. Since revenues and costs are increasing, as is driver aggressiveness, drivers face a moral hazard.

II. The Data

A. *Description of the TLC Data*

Data from the TLC contain information about all drivers of New York City yellow taxis from spring of 2005 and fall of 2007. The 2005 data contain driver-level records on three types of driving outcomes: accidents, convictions, and summonses. The accident records describe all taxi accidents with injuries or property damage exceeding $1,000 to any vehicle involved. The violations records describe all drivers' convictions for driving violations during taxi operation for which points are issued against a DMV driver's license. The accident and summons data originate from the New York City Police Department. The summons records describe all summonses issued for TLC violations (as opposed to DMV driving violations), such as refusing passenger service, using the taxi for unlawful purposes, and missing required items from the taxi, such as the driver's taxi-driving license on display.

The 2007 data only contain records for the summons outcomes. The other two outcomes originate from the New York City Police Department and are not available for 2007.[11] In addition to being available only for 2005, the accidents and convictions records also have more limited variation than the summons records. Approximately 24 percent of drivers receive at least one summons during a six-month period versus 14 percent having a conviction and 4 percent having an accident. Because of the greater variability, we focus our analysis on summonses. Nevertheless, we show that our results are robust for accidents and convictions.

[10] For example, some taxi owners indicated to us that brakes must be replaced every few weeks. Thus, whether drivers aggressively accelerate and brake, take care to avoid curbs, and so on can influence an owner's repair costs.

[11] We top-code outcomes to three to limit the impact of outliers, though less than 1 percent of drivers are affected.

The data also record each driver's country of birth, 2005 address, and, for drivers entering after 1997, an English-language test outcome and a written driving test score, both of which drivers must pass to obtain a taxi-driving license.[12] We calculate a driver's NYC taxi-driving experience from the number of months registered as an active NYC taxi driver, which we derive from his taxi driver's license numbers.[13]

Two driver-of-record files identify the taxi long-term lessee-driver and owner-driver was registered to operate on April 2, 2005, and October 21, 2007, which we use to merge the driver and owner data.[14] Owners receive a significant discount on insurance premiums when they register specific "named drivers" with the TLC and the insurance provider. These named drivers are then the exclusive operators of that taxi (including the owner himself when he drives), and these matches constitute the driver-of-record file.[15] Taxis with such named drivers represented 9,535 of the 12,779 taxis in 2005 and 9,025 of the 12,953 taxis in 2007. Since we know the driver-taxi match for these two years, we restrict analyses to the periods surrounding them.

For the long-term lessees, the driver-of-record files record the current lease period, which is typically a six-month interval. Specifically, 50 percent are between 177 and 183 days long, and 95 percent are between 147 and 189 days long. We match the dates of offense for the summonses, convictions, and accidents to these lease periods to identify the drivers' outcomes associated with particular taxis. To make the driving outcomes comparable across drivers with different lease period lengths, we use the recorded lengths to normalize the number of incidences to 182 days.

Additionally, we use census data on the fraction of residents in the driver/owner census tract from the same country of birth (resident country of birth is only reported at the tract level), which we match to drivers/owners using their addresses.[16] Table 1 reports summary statistics for the primary variables in our sample. On average, drivers have 11 years of NYC taxi-driving experience, 44 percent of drivers lease in-network, and approximately 1 in 3 drivers receives a summons, 1 in 6 a conviction, and one in 20 a serious accident per six-month period.

B. *Discussion of Driving Outcome Measures*

Our data do not contain measures of vehicle maintenance, repair, and replacement costs borne by the owner.[17] Instead, we examine three measures that either directly or indirectly reflect costs generated by the driver but are borne by the owner. One measure is accidents, which, unlike maintenance and repairs, can be attributed directly to individual drivers. The second measure is convictions for driving

[12] The driving test is on NYC geography and DMV and TLC regulations.

[13] Taxi-driving license numbers are issued sequentially, which allows us to identify when a license was issued.

[14] Taxis not in the driver-of-record file have "unspecified" drivers, and TLC officials and industry participants indicate they are nearly always short-term lessees. Since short-term lessees operate different taxis every day, they cannot be matched in the data to specific taxis or, hence, to owner countries and are excluded from the analysis.

[15] Because of the large financial incentives to accurately report named drivers, we believe these records are accurate. Furthermore, summonses for inaccurate named driver records are very infrequent.

[16] The country of birth recorded in the TLC data is occasionally more precise than the country of birth in the census data. Any additional countries are assigned to the broader census categories, such as "Other western Europe."

[17] These data would not reveal directly the costs generated by the lessee-driver since attributing repair costs to the owner-driver versus the lessee-driver would be challenging, given that both operate the taxi nearly every day.

Table 1—Summary Statistics

	2005				2007			
	N	Mean	SD	Min/Max	N	Mean	SD	Min/Max
Years of NYC taxi-driving experience	1824	11.0	7.5	0.5/33.0	1877	11.2	8.0	0.8/35.4
Lease in network	1824	0.45	0.50	0/1	1877	0.42	0.49	0/1
Number of summons	1824	0.39	0.76	0/3	1877	0.33	0.68	0/3
Number of accidents	1824	0.04	0.21	0/2				
Number of convictions for driving violations	1824	0.17	0.46	0/3				
Passed English language test on first try	731	0.90	0.30	0/1	957	0.92	0.28	0/1
Driving test score on first try	743	78.6	10.5	0/100	967	78.2	10.5	0/100
Fraction residents in driver's census tract from same country	1677	0.06	0.09	0.00/0.85				
Fraction residents in owner's census tract from same country	1678	0.08	0.15	0.00/0.94				

Notes: The table shows summary statistics of the sample of drivers used in the regression analyses. Fewer observations are available for the English-language and driving test results because these data are only available for drivers entering the industry after 1997 and for fraction residents in driver/owner census tract from the same country because of missing addresses. Accident, conviction, and address (and hence census) data are only available for 2005.

violations, such as speeding or disobeying other traffic laws, which are enforced and processed in the same way as ordinary motorists' driving violations. Both accidents and convictions directly reflect driver ability and aggressiveness and, hence, are likely to reflect average risk for owners from unsafe driving.

Our third measure is summonses, which are given for a range of TLC violations (as opposed to DMV driving violations). Some driver violations generate costs for owners directly; owners receive fines for missing items, missing trip records, and vehicle condition violations stemming from the actions of drivers to whom they lease. Some driver violations generate costs for owners indirectly. For example, using the taxi for an unlawful purpose may risk hazardous moving violations; shift violations, in particular smoking in the taxi; and even the taxi itself (i.e., temporary impounding). While not all violations affect owner costs directly, overall violations may still reflect general driving style and possibly a moral hazard via the level of vehicle usage since the number of violations likely increases with time on the road, and drivers may operate the taxi for more time within a lease period since they do not pay the full operating cost.

We believe some types of summonses better reflect a driver's driving style and conscientiousness than others. However, we have no way to identify precisely which summons types are most reflective; hence, we aggregate all summonses for the primary analysis. However, we also provide results broken out by summons type to show that individual types indeed conform to our expectations of which are more likely to be susceptible to moral hazard.

C. *Availability of Country-of-Birth Data*

In order to identify whether the driver is leasing from a same-country owner, our analysis requires the availability of country-of-birth data for both the lessee-driver and the owner from whom he leases. Table 2 reports the number of drivers with

Table 2—Availability of Data on Driver and Owner Country Of Birth

	2005				2007			
		Country data available				Country data available		
			In network	Out of network			In network	Out of network
All long-term lessees (N)	13,523	1,955	874	1,081	12,683	2,053	850	1,203
Mean experience (years)	9.3	10.6	11.1	10.2	9.6	10.7	11.8	9.9
Drive owner-driver taxis (N)	2,540	1,504	782	722	2,548	1,604	770	834
Mean experience (years)	10.3	10.9	11.1	10.6	10.7	10.9	11.8	10.0
Drive non-owner-driver taxis (N)	10,983	451	92	359	10,135	449	80	369
Mean experience (years)	9.1	9.9	11.3	9.5	9.3	10.2	11.9	9.8

Note: The table shows number of drivers for whom country-of-birth data are available for both the driver and the corresponding owner, by ownership arrangement, along with mean experience levels.

available country data for both the driver and owner, along with the mean NYC taxi driving experience of these drivers.[18] Country data for both the lessee and owner are available for 1,955 and 2,053 lessees in the 2005 and 2007 data, respectively. For driver experience and all other observable characteristics (e.g., race, English language test, driving test [not shown]), those with data are similar to those without, suggesting that the drivers we analyze are similar to the drivers for whom country data are missing.[19]

Most lessee-drivers with available country data for both the driver and the owner share a taxi with an owner-driver (i.e., they operate owner-driver taxis).[20] Thus, the sample we analyze consists primarily of lessee-drivers who lease from owner-drivers. This sample is ideal for studying social networks because owner-drivers are likely to be invested in the condition of their vehicle since they drive most days and to interact with the lessee regularly, thereby facilitating a social connection. Nondriving owners, who often own hundreds of taxis, may have more anonymous relationships with drivers. Table 2 shows that a small fraction of lessee-drivers who have country-level data available lease from nondriving owners. The results are similar and somewhat stronger when they are excluded.

D. *Prevalence of In-Network Leasing*

In our data, 44 percent of drivers are in-network,[21] suggesting that there are benefits to in-network driving. However, it is possible that a "proximity" effect, based on residential clustering by country of birth, rather than an "in-network" effect may be driving this result. Therefore, we test this hypothesis directly. If the countries of

[18] Table 1 reports slightly fewer observations than Table 2 because Table 1 describes only observations for which country-level data are available, and data are nonmissing for all other variables in the regression analysis, while Table 2 only describes observations for which country data are available regardless of the availability of other variables.

[19] Fifty-four, 23, and 13 percent of drivers with country data were Asian (including South Asian), black, and white, respectively, while 58, 22, and 8 percent of drivers without country data were Asian, black, and white, respectively.

[20] Country data are available primarily for lessee-drivers on owner-driver taxis only because country data are available for most drivers, including owner-drivers, but not for non-driving owners.

[21] See Table 9 in the online Appendix for the distribution of driver countries for the largest owner countries.

birth of the owners and drivers do not determine the owner-lessee match, then, conditional on having any owners in the same census block, the owner's country should not affect the likelihood of leasing from an owner who lives in the same block. In fact, the likelihood of leasing from a neighbor is 32 percentage points higher ($p < 0.01$) when there is an owner from the same country in the block (conditional on having any owner). Also, conditional on having an owner from the same country in the same block, having any owner from any other country in the block has a much smaller (though still significant) effect on leasing from that owner. This result suggests that while proximity is certainly an important factor, the prevalence of in-network driving cannot be explained by proximity alone, indicating real benefits to leasing in-network.

III. Empirical Strategy

In this study, we aim to remove any selection bias to uncover the causal (effort) effect of driving a taxi owned by someone from their same country and document the degree and direction of any selection on ability into in-network driving. To identify the causal effect of in-network driving, we use two distinct sources of arguably exogenous variation in in-network driving. Results from these distinct strategies offer robustness checks on each other.

A. *Baseline Model*

Our basic empirical strategy is to compare the driving outcomes of drivers who lease from an owner from the same country of birth to the driving outcomes of drivers who lease from an owner from a different country of birth. Specifically, we estimate the following model by OLS:

$$Y_{it} = \beta \cdot Same_{it} + \gamma \cdot \ln(Exp_{it}) + \alpha_{\mathrm{i}} + \varepsilon_{it}. \tag{1}$$

In (1), Y_{it} is the outcome of driver i at time t, which is the number of accidents, convictions, or summonses; $Same_{it}$ is an indicator variable denoting whether driver i leases from an owner from the same country of birth at time t; Exp_{it} is the experience level of driver i at time t; α_{i} is the unobserved (to the researcher) time invariant ability level of driver i; and ε_{it} is the idiosyncratic error term. Because driver ability is not observed, the total error term is $\alpha_i + \varepsilon_{it}$.

OLS estimates of β from (1) are unlikely to yield the causal effect of in-network driving because owners may select the drivers to which they lease based on unobserved driver ability. In the presence of screening, where owners can better observe the ability of same-country drivers, such that same-country drivers possess higher ability, there would be positive selection, and the OLS estimates of β would be biased downward (i.e., more negative). Alternatively, if owners know that same-country drivers will exert more effort, making owners more willing to hire same-country drivers even if they are of lower ability, there would be negative selection, and the OLS estimates of β would be biased upward (i.e., less negative). As such, comparing the OLS estimates to estimates obtained from models that isolate the causal effect of being a same-country driver are informative about the selection process. We describe two strategies to remove the effects of selection below.

B. *Within-Driver Identification Strategy*

The first strategy for isolating the effort effect is to use within-driver variation in same versus different-country leasing. Specifically, with two observations per driver over time, in 2005 and 2007, we can compare the outcomes of drivers who leased from a same-country owner in one period and a different-country owner in another period. By comparing the outcomes of the same drivers over time, we effectively remove the contribution of time-invariant ability on driving outcomes (i.e., the contribution due to selection), thus isolating the effort effect. Specifically, we estimate the following differenced (within-driver) model by OLS:

$$\Delta Y_i = \beta \cdot \Delta Same_i + \gamma \cdot \Delta \ln(Exp_i) + \Delta \varepsilon_i. \tag{2}$$

All variables are defined as before, but now, the time-invariant ability component is differenced away so that the error term $\Delta \varepsilon_i$ does not reflect the effects of selection. Identification of the effort effect β comes from drivers who switch from leasing from a same-country owner to leasing from a different-country owner (and vice-versa). The identifying assumption is that any changes in unobserved characteristics that affect driving outcomes are orthogonal to changes in same-country driving status over time. It is worth noting that this model estimates a local treatment effect for "switchers," who make up only 3.2 percent of the sample (58 drivers); therefore, while the effect may be well identified, it does not necessarily generalize to all drivers. This is one motivation for our use of another completely different source of variation to identify effort effects.[22]

C. *Cross-Sectional Instrumental Variables Strategy*

In our second approach, we exploit the variation in leasing from a same-country owner that is due to the residential locations of drivers versus same-country owners. The idea is that drivers may be more likely to lease from an owner who lives nearby rather than far away, everything else being equal. Thus, a driver who happens to live near a same-country owner should be more likely to lease from that owner simply because he lives close to that owner. Using the 2005 addresses of drivers and owners, we compute the minimum distance of each driver to a same-country owner. We then use this distance as an exogenous instrument for in-network leasing to isolate the causal effect of in-network driving.

[22] Nevertheless, switchers and nonswitchers are observably similar. On average, switchers have 10.9 years of experience and are 44.4 years old, whereas nonswitchers have 11.1 years of experience and are 44.5 years old. Sixteen, 12, 11, 9, and 10 switchers are from Bangladesh, Pakistan, India, Haiti, and other countries, respectively, which is proportionally similar to nonswitcher countries of birth.

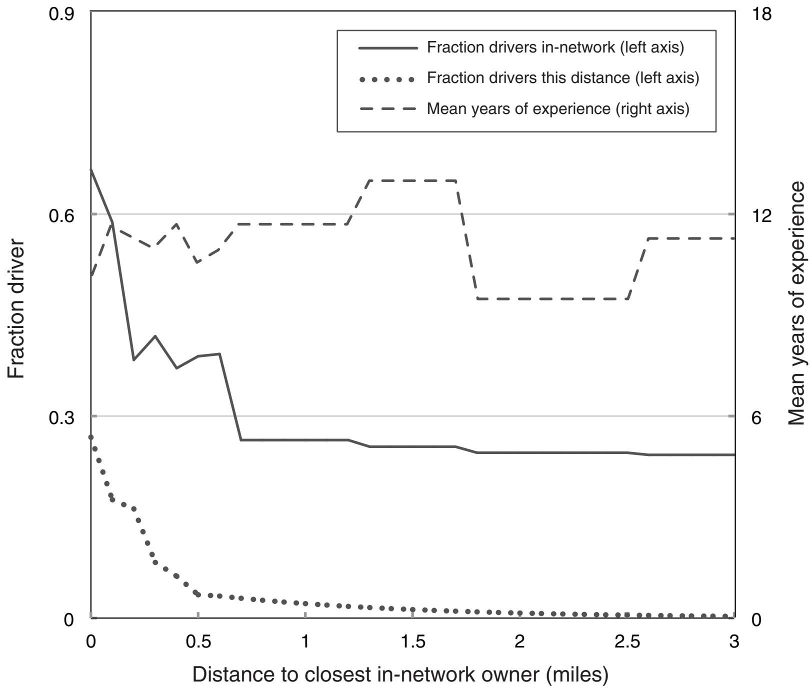

Figure 1. Graphical Evidence in Support of the Validity of the Instrument

Note: From top to bottom, the curves represent mean experience, fraction of drivers leasing in-network, and fraction of drivers at that distance, all versus distance to an in-network owner.

To be a valid instrument, this distance measure must (a) be correlated with in-network driving, (b) not be correlated with unobserved ability, and (c) have a monotonic relationship with being a same-country driver. Figure 1 shows clearly that distance to the closest same-country owner predicts leasing from a same-country owner. The effect, however, is most pronounced for distances of less than two-tenths of a mile, and for this reason, we specify our instrument $Dist_i$ as an indicator of whether the driver lives in the same census block as a same-country owner.[23] Figure 1 also shows that distance to a same-country owner is unrelated to driving experience, the best measure of driver ability in our data. It is also clear from Figure 1 that the relationship between distance to a same-country owner and being a same-country driver is monotonic. We present formal econometric tests of the validity of the instrument in Section IV.[24]

[23] Using the same census tract or log of distance to the nearest same-country owner gives similar though modestly weaker results due to a weaker first stage.

[24] A slightly different interpretation of the instrumental variable estimate is that leasing from a nearby same-country owner facilitates stronger network effects because an owner can better enforce social sanctions. However, the OLS estimates are similar for drivers living nearby versus far away from the owner, indicating that network effects are not predicated on proximity. Regardless, the alternate interpretation would still indicate the network effects.

In summary, we exploit this variation in drivers' residential locations vis-à-vis owners to predict network status and then use these predicted values in a regression on driving outcomes. Specifically, we estimate the following system of equations by 2SLS:

$$Same_i = \delta_1 \cdot Dist_i + \delta_2 \cdot \ln(Exp_i) + \theta_{1c} + \theta_{1co} + v_{1i} \tag{3}$$

$$Y_i = \beta \cdot Same_i + \gamma \cdot \ln(Exp_i) + \theta_{2c} + \theta_{2co} + v_{2i}. \tag{4}$$

In (3) and (4), all variables are defined as before, but there is no longer any time variation. θ_c is a driver country-of-birth fixed effect, and the error term ν includes both idiosyncratic variation and driver ability. With driver country-of-birth fixed effects, this model exploits variation in residential locations of drivers *who come from the same country of birth* to remove selection to in-network driving. We also include fixed effects for the *owners'* country of birth, θ_{co}, to account for any systematic differences across the owners' places of birth.

IV. Results

A. *Baseline Analysis*

Table 3 provides regression evidence regarding the effect of in-network driving on driving outcomes. Column 1 shows OLS results for a driver's number of summonses. Drivers leasing in-network have 0.09 fewer summonses per six-month period ($p < 0.01$), which is a modest-size effect relative to the standard deviation of the number of summonses per driver of 0.72. Both driver and owner-country fixed effects are included in column 2 to address the possibility that drivers or owners from some countries are systematically better or worse in ways that are correlated with the probability of in-network leasing, and the results are similar.[25,26]

B. *Within-Driver Analysis*

Column 3 exploits the panel nature of the data to include driver fixed effects to remove any bias due to selection across countries into in-network driving and selection within countries but across drivers into in-network driving. The within-driver estimate indicates that drivers leasing in-network have 0.334 fewer summonses per six-month period ($p < 0.01$), substantially larger than the OLS estimate.

As a check of the identifying assumption that changes in in-network status between 2005 and 2007 are not associated with unobserved changes in factors that

[25] Results are similar with Poisson and negative binomial regression models and are available in an online Appendix.

[26] All specifications, including driver age and/or census neighborhood characteristics, such as median income and education level, have, at most, a small effect on outcomes after controlling for experience and no effect on the estimate of in-network effect. We, therefore, use the more parsimonious specification that includes only experience.

Table 3—Models of Summonses

	OLS	OLS	Within driver	Within driver	2SLS	2SLS	Country level
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Lease in-network	−0.085***	−0.070**	−0.334***	−0.489***	−0.251**	−0.333	−0.350**
	[0.024]	[0.033]	[0.093]	[0.151]	[0.128]	[0.389]	[0.138]
Lease in-network × lease out of network 2005				0.271			
				[0.204]			
Log years of experience	−0.105***	−0.108***	−0.315*	−0.322*	−0.082***	−0.086***	−0.058
	[0.014]	[0.016]	[0.167]	[0.167]	[0.022]	[0.025]	[0.092]
2007 indicator	−0.064***	−0.054**	−0.015	−0.026			−0.225***
	[0.023]	[0.024]	[0.057]	[0.058]			[0.081]
Owner-driver summonses for lessee country							−0.143
							[0.209]
Constant	0.648***	1.162***	1.248***	0.665***	1.042	0.785***	0.636***
	[0.038]	[0.329]	[0.339]	[0.073]	[0.772]	[0.200]	[0.078]
First stage coefficient on lease in-network					0.312	0.115	
First stage F-statistic					128.8	29.9	
Fixed effects for driver country	No	Yes	No	No	No	Yes	No
Fixed effects for owner country	No	Yes	No	No	No	Yes	No
Fixed effects for driver	No	No	Yes	Yes	No	No	No
Observations	3,646	3,646	3,646	3,646	1,629	1,629	149
R^2	0.025	0.091	0.039	0.041	0.007	0.110	0.086

Notes: The dependent variable is number of summonses per driver. The 2SLS sample is limited to 2005 since addresses are only available for this year. Heteroskedasticity robust standard errors that are adjusted for clustering at the country-of-birth level are reported in brackets.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

may affect driving outcomes, we break out the effect separately for the 25 drivers switching from in-network to out-of-network leasing between 2005 and 2007 versus the 33 drivers switching from out-of-network to in-network leasing between 2005 and 2007. If switching is endogenous (i.e., if drivers tend to switch after an accident but then have good outcomes), then the effects of switching into versus out of network might be different. Column 4 breaks out the effect for drivers switching away from in-network driving between 2005 and 2007 ("lease in-network") versus switching into in-network leasing between 2005 and 2007 ("lease in-network × lease out of network 2005," which is additive to the "lease in-network" estimate). While drivers switching away from in-network driving show a larger effect, the estimates for both types of switchers are much larger than the OLS estimates, and the two types are not statistically different ($p = 0.18$).[27]

[27] Note that even though all drivers in the panel had two additional years of experience in 2007 versus 2005, the effect of log of experience is identified due to the nonlinearity of the measure. The nonlinear measure of experience is suitable given that the marginal benefit of experience almost certainly decreases with experience.

Table 4—Falsification Tests for the Excluded Instrument

	Log years of experience	Log years of experience	Passed English test	Passed English test	Driving test score	Driving test score	Driver effect from panel	Driver effect from panel
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
In-network owner in block	−0.088	−0.089	−0.012	−0.004	0.86	1.177	0.010	−0.018
	[0.056]	[0.060]	[0.027]	[0.034]	[0.987]	[1.143]	[0.039]	[0.045]
Constant	2.075***	1.414***	0.910***	0.941***	78.4***	87.6***	0.018	0.350
	[0.026]	[0.509]	[0.013]	[0.029]	[0.468]	[2.486]	[0.022]	[0.559]
Fixed effects for driver country	No	Yes	No	Yes	No	Yes	No	Yes
Fixed effects for owner country	No	Yes	No	Yes	No	Yes	No	Yes
Observations	1,629	1,629	648	648	659	659	1,629	1,629
R^2	0.002	0.191	0.000	0.161	0.001	0.168	0.000	0.131

Notes: The dependent variables are measures of driver ability (labeled at column head). The explanatory variable is the excluded instrument. The sample is limited to 2005 since addresses are only available for this year. Heteroskedasticity robust standard errors that are adjusted for clustering at the country-of-birth level are reported in brackets.

*** Significant at the 1 percent level.

C. *Instrumental-Variables Analysis*

Columns 5 and 6 contain the instrumental variables results, where the instrument for in-network driving is an indicator of whether the driver lives in the same Census block as an owner from the same country. The first-stage F-statistic for the excluded instrument is 131 and 41, depending on whether driver and owner country-fixed effects are included.

The instrumental variables estimate without country fixed effects indicates that drivers leasing in-network have 0.251 fewer summonses per six-month period ($p < 0.05$). Again, the estimate is substantially larger than the OLS estimates ($p = 0.05$). When driver and owner-country fixed effects are included, the estimate is very similar though less precise, given the weaker first stage.[28,29]

In Table 4, we provide evidence from falsification tests that the instrument is uncorrelated with driver ability by regressing our instrument on proxies for driver ability. Columns 1 and 2 show that the instrument has approximately zero effect on experience. Columns 3–6 show that the instrument also has essentially no effect on whether the driver passed an English language test or on the driver's score on a driving test. Finally, we examine a direct measure of unobserved driver ability. Specifically, because we have longitudinal data while our instrument relies on cross-sectional variation, we can obtain estimates of driver ability (α_i) from equation (1) using a fixed effects model and then see if these estimates are correlated with the

[28] Columns 6 and 7 also report the estimated coefficients on the excluded instrument in the first-stage regression, showing the instrument to be a very strong predictor of network status.

[29] Note that the sample size for these specifications is smaller than previous specifications since the instrument is based on driver and owner addresses, which are only available in the 2005 data.

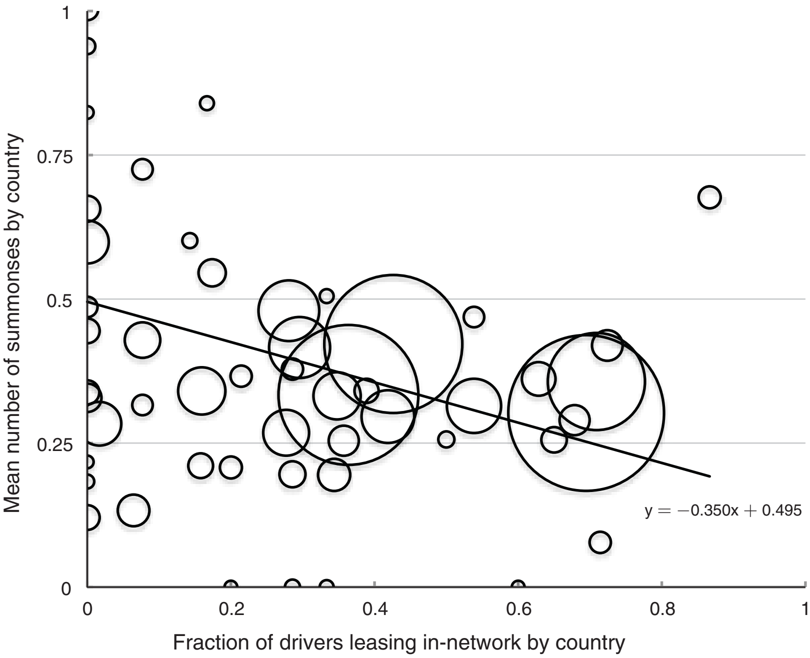

Figure 2. Country-level Relationship between In-network Leasing and Summons

Notes: Circle size represents the number of drivers from that country. The line is the fitted relationships between the fraction of drivers leasing in-network by country versus the mean number of summonses by country, unweighted by number of drivers by country (the weighted fit is similar).

instrument.[30] Columns 7 and 8 show that the relationship between ability and our instrument is approximately zero.

V. Robustness Checks

In this section, we provide additional evidence that our estimates represent causal effects. Specifically, we present estimates from a third independent source of variation. We show that the effects are strongest for summons types most susceptible to moral hazard, and we provide evidence that our results are not driven by sample selection or selection on the part of owners.

A. *Country-Level Analysis*

Our identification strategies above exploit variation in network status within country-of-birth networks. To provide additional evidence that our results are not driven by selection within a network, we also aggregate the data to the country level and use only cross-country variations.

[30] We also use a mixed effects estimator to obtain the Best Linear Unbiased Predictors (BLUPs) of ability and find these are also not correlated with our instrument. This method for identifying BLUPs is used in Jackson (2010).

We begin by showing in Figure 2 the relationship between the mean numbers of summonses per driver by country versus the fraction of drivers from that country leasing from an owner from the same country.[31] The figure shows that driving outcomes are better for drivers from countries with more in-network leasing. We investigate the relationship more formally by estimating equation (1) aggregated at the country-of-birth level. To address the concern that in-network leasing may be correlated with average driver ability across countries, we control for the average number of summonses of *owners* from the same country of birth as the lessee.

Column 7 of Table 3 provides these results, showing that the in-network leasing effect is substantially larger than in columns 1 and 2, and is statistically significant ($p = 0.01$). The estimate indicates that a country with all in-network driving will have 0.35 fewer summonses per driver than a country with no in-network driving. Because the cross-country estimates could be subject to bias from selection *across* networks but not from selection *within* networks, the fact that the cross-country estimates are larger than the OLS and are similar to the within-driver and 2SLS results suggests that much of the selection occurs within country-of-birth networks, there is negative selection into in-network driving, and both proposed methods to remove selection are likely valid.

B. *Analysis by Summons Type and Alternative Driving Outcomes*

If effort effects are indeed driving the observed network effects, we would expect network effects to be largest for the types of summonses that directly impact owners' costs. For example, we expect to find a relatively large effect for hazardous driving summonses (since they may reflect more reckless driving and, hence, higher associated costs from accidents and mechanical defects for the owner) and for trip sheet, vehicle condition, and missing items summonses (since owners receive a fine when drivers to whom they lease incur these summonses).

These results are more speculative since we cannot identify precisely the level of moral hazard associated with each summons type and because of the reduced test power due to splitting the data more finely. Nevertheless, the results in Table 5, which reports both OLS and instrumental variables estimates, are generally consistent with effort effects.[32] In-network driving has the most significant effect on the three summons types that directly result in an owner fine: trip record violations, vehicle condition violations, and missing item violations. While the *t*-statistics for hazardous moving violations and for using the taxi for unlawful purposes, which may also directly generate costs for the owner, are smaller, they have the correct sign, and the instrumental variables estimates are among the most negative.[33]

For the reasons discussed in Section II, we focus our analysis on summonses. However, we also examine accidents and convictions for driving violations as alternative driving outcomes, albeit with less power and without a panel data component,

[31] Countries with at least 5 drivers are included (55 countries), and the circle area is proportional to the number of drivers. Country labels are not included due to a TLC request that country-level driving outcomes not be reported.

[32] We do not include country or driver fixed effects due to insufficient variation at the summons level.

[33] Instead of coefficients, *t*-statistics are reported since number of summonses varies by type and is not comparable.

Table 5—Results by Individual Summons Type

Summons for TLC driver violation	N OLS (2005, 2007)	*t*-stat OLS (2005, 2007)	N 2SLS (2005)	*t*-stat 2SLS (2005)	Corresponding TLC owner violation or risk
Non-hazardous moving violation	169	0.95	84	0.28	
Stationary vehicle violation	104	0.00	30	−0.06	
Shift violation (primarily cell phone and smoking)	79	−0.37	35	−2.39	
Abuse (primarily verbal abuse)	64	−0.41	39	−0.10	
Passenger refusal	118	−0.43	43	−0.26	
Hazardous moving violation	95	−0.44	60	−1.99	Risk to taxi
Passenger request violation	61	−0.69	20	1.10	
Noncooperation with TLC	104	−0.72	74	−0.12	
Using taxi for unlawful purpose	211	−0.98	96	−1.32	Risk to taxi
Off-duty procedures violation	56	−1.02	36	−0.81	
Discourteous to passenger	85	−1.32	44	0.03	
Trip records violation	153	−1.70	90	−2.12	Missing driver records violation
Vehicle condition violation (primarily safety related)	84	−2.06	34	0.36	Vehicle condition violation
Item missing from taxi (primarily trip records)	184	−3.42	97	−1.82	Item missing from taxi (primarily trip records)

Notes: The table reports number of drivers with a positive number of that summons type (N OLS, N 2SLS) and the *t*-statistic for regressions of in-network driving on the probability the driver incurred the summons type listed in the left-most column. Results for OLS (2005 and 2007) and 2SLS regressions (2005 only) are listed. Fixed effects for driver and owner country are not included (to preserve test power). The corresponding direct risk/cost to the owner is listed in the right-most column.

since only 2005 outcomes are available. Accident rate in particular is an attractive measure because it captures an outcome that affects owners directly. In columns 1–4 and 5–8 of Table 6, the dependent variables are number of accidents and number of convictions for driving violations, respectively, by the driver during the lease period. We report baseline OLS estimates, OLS estimates with fixed effects for driver country, country-level estimates, and instrumental variables estimates. Reassuringly, all of the estimates have the correct sign and a reasonable magnitude. While imprecise, the country-level and instrumental variables estimates, which are likely to isolate the causal effect of in-network driving on effort, are again larger than the OLS estimates, indicating an important network effect and that in-network drivers have lower ability on average than out-of-network drivers.

C. *Discussion of Sample Selection*

Given that we only have data on individuals who became drivers and not on prospective drivers not already in the industry, readers may wonder whether our findings are driven by some sample selection bias (as distinct from selection into in-network driving conditional on being in the sample). Specifically, one concern might be that high-quality prospective drivers are more likely to enter the industry than in-network drivers, while low-quality prospective drivers do not enter the industry at all. Such sample selection could potentially bias the driver-level cross-sectional results (OLS and 2SLS). However, this selection would not drive the within-driver results because all comparisons are within the same driver over time. Also, it would not drive the aggregate country-level results as long as the sample selection process

Table 6—Models of Alternative Driving Outcomes

	Accidents				Convictions			
	OLS	OLS	Country level	2SLS	OLS	OLS	Country level	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lease in-network	−0.022**	−0.019	−0.095*	−0.027	−0.035*	−0.023	−0.068	−0.057
	[0.011]	[0.015]	[0.051]	[0.043]	[0.021]	[0.035]	[0.103]	[0.079]
Log years of experience	−0.016**	−0.012	−0.081**	−0.018**	−0.017	−0.055***	−0.104**	−0.017
	[0.006]	[0.007]	[0.036]	[0.008]	[0.012]	[0.014]	[0.049]	[0.014]
Owner-driver outcomes for lessee country			−0.148**				−0.125*	
			[0.070]				[0.073]	
Constant	0.067***	0.071**	0.225**	0.074***	0.221***	0.351**	0.411***	0.224***
	[0.014]	[0.034]	[0.086]	[0.018]	[0.030]	[0.156]	[0.124]	[0.040]
Fixed effects for driver country	No	Yes	No	No	No	Yes	No	No
Fixed effects for owner country	No	Yes	No	No	No	Yes	No	No
Observations	1,813	1,813	79	1,647	1,855	1,855	81	1,687
R^2	0.006	0.094	0.145	0.009	0.014	0.114	0.078	0.005

Notes: The dependent variable in columns 1–4 and 5–8 are driver-level number of accidents and driving convictions, respectively. The sample is limited to 2005 since outcome data are only available for 2005. Heteroskedasticity robust standard errors adjusted for clustering at the country level are reported in brackets.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

was the same across countries because this selection within country-of-birth will be averaged out at the country level. Given the robustness of our central findings across models that are, and are not, susceptible to sample selection bias, the results do not appear to be driven by sample selection. Moreover, all of our strategies that account for selection (conditional on being in sample) indicate that the selection into in-network driving is negative rather than positive, which is generally inconsistent with sample selection driving our results.

D. *Discussion of Owner Selection*

Even though our results suggest that selection on the part of drivers does not drive our results, one may wonder about selection on the part of owners. Specifically, if owners who lease to an in-network driver are more risk averse than owners who lease to an out-of-network driver, then the owners leasing in-network may systematically use contracts with stronger incentives (i.e., requiring larger deposits) for better driving outcomes. If so, our results might reflect owner selection rather than a mitigation of moral hazard.

This possibility is unlikely since, as discussed in Section I, owners have only modest leeway to vary contracts, given the regulatory caps on the deposit amount ($500) and lease fee ($666), which industry participants indicate are respected in practice. That is, there is little flexibility in contracting terms for explaining the large observed difference in outcomes between in-network and out-of-network drivers. However, we examine this issue more formally.

Table 7—Robustness Checks

	Lessee-driver					Owner-driver	
	Within owner (1)	IV (2)	Country level (3)	Lessee-only taxis (4)	Lessee-only taxis (5)	Within owner (6)	IV owner (7)
Lease in-network	−0.091	−0.140	−0.209*	−0.113*	−0.132	0.090	0.074
	[0.073]	[0.114]	[0.121]	[0.062]	[0.099]	[0.076]	[0.141]
Log years of experience	−0.066***	−0.086***	−0.073	−0.114***	−0.091***	−0.335	−0.107*
	[0.024]	[0.025]	[0.085]	[0.030]	[0.035]	[0.492]	[0.058]
2007 indicator	−0.092***		−0.206***	−0.077	−0.055	−0.030	
	[0.030]		[0.072]	[0.052]	[0.056]	[0.082]	
Owner-driver summons rate for lessee country			−0.203				
			[0.308]				
Constant	−0.059	0.651***	0.733***	0.659***	3.291***	0.358	0.447**
	[0.832]	[0.063]	[0.257]	[0.078]	[0.085]	[1.229]	[0.220]
Fixed effects for driver country	Yes	No	No	No	Yes	No	No
Fixed effects for owner country	No	No	No	No	Yes	No	No
Fixed effects for owners	Yes	No	No	No	No	Yes	No
Observations	3,646	1,615	156	769	769	1,787	795
R^2	0.069	0.009	0.071	0.028	0.214	0.021	0.015

Notes: The dependent variable in columns 1–5 is the number of summonses incurred by the lessee. In column 2, the instrument is an indicator for the owner living in the same census block as any same-country driver. In column 3, "lease in-network" is the fraction of owners for the country leasing to same-country drivers. The sample in columns 4–5 are lessees of taxis operated exclusively by lessees (i.e., owned by a non-driving owner). In column 6–7, the dependent variable is the number of summonses incurred by the owner-driver, and "lease in-network" indicates whether the lessee-driver to whom the owner-driver leases is in-network. Column 6 includes owner fixed effects. In column 7, the instrument is an indicator for the owner living in the same census block as any same-country lessee.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

In Table 7, we present empirical evidence that driver selection does not drive our results. First, we estimate our main OLS model with owner fixed effects instead of driver fixed effects to compare outcomes of in-network drivers and out-of-network drivers with the same owner. The results are almost identical to the basic OLS model, showing that behaviors specific to the owners (e.g., risk preferences or the types of contracts they offer drivers) do not drive our results. Second, we run instrumental variables models, using an indicator for whether there is a same-country driver living in the same census block as the owner as an instrument for whether the owner leases in-network. The results again indicate that owner selection is not driving the results and that a causal network effect is present. Third, we present the country-level lessee driving outcomes versus the fraction of *owners* from that country who lease in-network. These results are similar to the original country-level model and, again, are consistent with a network effect and not owner selection. We also run the basic OLS results when we restrict the sample to lessees who lease from owners who are not themselves drivers and show that the results are essentially the same.[34] As a final check on the possibility of owner selection, we show the instrumental variables

[34] For all of these results, the magnitudes of the estimates are smaller than the estimates presented earlier that control for driver selection. This is as expected, given that the models in Table 7 do not directly control for driver selection. (There is insufficient variation in the data to control for both driver and owner selection in the same model.)

and within-owner models using owner-driver driving outcomes and the instrument based on owner residential location, where "lease in-network" indicates that the owner leases to an in-network driver. For both of these models, the point estimates are small relative to those found for lessees, and the null hypothesis that in-network owners are the same as out-of-network owners cannot be rejected. In sum, the earlier results of an important in-network effect are robust, explicitly including various types of controls for owner selection.

VI. Exploration of Mechanisms

Social network effects can operate through at least five mechanisms, including screening, monitoring costs, social sanctions, social preferences, and social collateral. In this section, we discuss which of these mechanisms is most consistent with our findings, and we present suggestive evidence on the mechanisms.

First, we consider screening and monitoring costs. Since the OLS estimates show a significantly smaller benefit of in-network driving versus the within-driver, instrumental variables, and country-level estimates, we have evidence that selection is such that in-network drivers having lower ability, on average, than out-of-network drivers. Under screening, there should be selection of higher ability drivers into in-network leasing. Thus, screening is unlikely to explain the driver-owner matches and outcomes.

Second, monitoring costs are unlikely to explain the network effects for several reasons. First, taxis are operated primarily in Manhattan, typically far from owners and members of the community, who are primarily located in Brooklyn and Queens; hence, direct monitoring of lessees is not feasible at reasonable cost, regardless of network status.[35] Second, most drivers in the sample lease from an owner who himself drives the taxi, giving any owner some ability to monitor vehicle condition, regardless of network status.

Third, we examine the effect of network density on driving outcomes to learn about social sanctions, social collateral, and social preferences. Following Bertrand, Luttmer, and Mullainathan (2000) and Gil and Hartmann (forthcoming), we use census data on the density of residents from a particular country who live in a neighborhood to measure network strength. Specifically, we measure driver/owner network strength as the fraction of residents in the driver/owner census tract from the same country of birth. While this measure is imperfect, it nonetheless should reflect the level of communication and ties between residents, for example, through the number of community organizations and events. The idea is to test whether denser owner/driver networks generate larger improvements in outcomes, which may suggest the mechanisms at play.[36]

We also consider social sanctions, which would appear as community-enforced punishments against an offending group member. These sanctions would likely

[35] One might argue that other drivers could observe a driver's driving style and report this information to the owner. However, this possibility seems remote due to the small chance that, out of the approximately 40,000 drivers in the industry (short-term lessees, long-term lessees, and owner-drivers), any single driver observes another driver he knows exhibiting poor driving, also knows the owner, and wants to report the observed driving style to that owner.

[36] We examine 2005 outcomes only, since address data, which identify tracts, are available for 2005 only.

Table 8—Models of the Effect of Network Strength

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lease in-network	−0.091**	−0.080**	−0.075*	0.055	−0.106***	−0.092**	−0.096**	0.025
	[0.041]	[0.040]	[0.043]	[0.061]	[0.037]	[0.038]	[0.039]	[0.055]
Lease in-network × driver network strength	−0.239		0.077	−0.071	−0.014		0.074	0.057
	[0.230]		[0.326]	[0.371]	[0.091]		[0.109]	[0.129]
Lease in-network × owner network strength		−0.386*	−0.603*	−0.573		−0.152**	−0.217***	−0.220**
		[0.200]	[0.323]	[0.376]		[0.068]	[0.077]	[0.094]
Log years of experience	−0.094***	−0.091***	−0.087***	−0.089***	−0.095***	−0.091***	−0.087***	−0.089***
	[0.022]	[0.022]	[0.023]	[0.026]	[0.022]	[0.022]	[0.023]	[0.026]
Constant	0.629***	0.622***	0.612***	0.867	0.631***	0.622***	0.612***	0.885
	[0.055]	[0.055]	[0.057]	[0.734]	[0.055]	[0.055]	[0.057]	[0.743]
Fixed effects for driver country	No	No	No	Yes	No	No	No	Yes
Fixed effects for owner country	No	No	No	Yes	No	No	No	Yes
Observations	1,644	1,645	1,518	1,518	1,644	1,645	1,518	1,518
R^2	0.020	0.020	0.019	0.158	0.020	0.020	0.020	0.158

Notes: The dependent variable is the driver-level number of summonses. Network strength in columns 1–4 is fraction of residents in the driver/owner census tract from the same country and in columns 5–8 is an indicator for at least 15 percent of residents from the same country. Heteroskedasticity robust standard errors adjusted for clustering at the country level are reported in brackets.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

operate via the owner's network (and not a driver's network); hence, a relatively dense owner network may facilitate these sanctions. Social collateral could operate similarly to social sanctions in that owners may provide or withhold favors to drivers based on driving outcomes, and these favors could operate through the owner's network. Therefore, the effect of social sanctions and social collateral should be reflected in an owner network strength effect. In contrast, social preferences would likely operate independently of an owner's network (though perhaps it would operate via the driver's network if this represents the driver's level of connection to the community).

In Table 8, we report regression results where the dependent variable is the number of summonses and the key explanatory variables are indicators for in-network leasing and the interaction of in-network leasing and network strength (i.e., an indicator variable denoting whether more than 15 percent of residents in the census tract are from the same place of birth).[37] Columns 5, 6, and 7–8 report results for driver network strength, owner network strength, and both, respectively. The results indicate that owner networks are more important than driver networks, and results can be seen most clearly in column 8, where driver network strength has no effect on summonses after conditioning on owner network strength. Also, the stand alone

[37] Columns 1–4 report similar results using the percentage of residents in the census tract instead of the nonlinear specification described.

in-network effect approaches zero, suggesting that the effects rely on a dense owner network.[38,39]

The estimates indicate that network effects operate via owner networks and not driver networks, and that network effects are weak when owner networks are weak. While these results are not dispositive, they suggest that social sanctions and/or social collateral are more important than social preferences for enabling the observed network effects.[40]

VII. Concluding Remarks

Moral hazard plays a central role in the understanding of numerous contractual relationships, and researchers have found evidence consistent with moral hazard in many contexts. However, much less is known about how to mitigate moral hazard problems when contracts are incomplete. While some evidence shows that social ties can mitigate moral hazard in group lending schemes in developing countries, little is known about the role of social ties in mitigating moral hazard outside of that particular context. In this study, we compare the driving outcomes of NYC taxi drivers who lease from owners from their same country to those of drivers who lease from owners from a different country. Across several models that exploit distinct and plausibly exogenous sources of variation to account for selection, drivers have better outcomes when driving in-network versus out-of-network, indicating a real causal effort effect. We also find that the marginal effects of driving in-network depend on the owner's network but not the driver's network, which suggests an important role for social sanctions.

An interesting policy implication from the group lending literature is the possibility of joint liability programs among groups of taxi drivers from the same network when the number of in-network owners from whom to lease is insufficient. Given that responsibility for mechanical defects typically cannot be contractually assigned to any single driver and that the costs of accidents often cannot be recovered from individual drivers, group liability may induce a higher level of effort among these drivers and allow more flexibility for owners and drivers to optimally form business relationships.

More generally, our results indicate that social ties can have incentive benefits in addition to positive selection effects. In many labor markets, jobs are frequently obtained through referrals via social ties. It is commonly thought that the main benefit of these referrals is to overcome asymmetric information over candidate quality. However, our results suggest these referrals may also help to overcome the moral hazard problem. More generally, our findings provide direct evidence that even in developed nations with strong institutions, social connections can play a valuable role in improving economic outcomes.

[38] The number of observations varies slightly across specifications due to differences in the availability of addresses.

[39] In columns 5–8, the continuous measure of network strength is replaced with an indicator for whether at least 15 percent of residents in the owner/driver census tract have the same country of birth. The idea is that a minimum density of residents might be necessary for a mechanism that operates through a local community to be effective, but that increasingly dense networks have marginally smaller benefits. Indeed, the results are more pronounced. We choose 15 percent since the marginal effect of network density appears to begin leveling off here.

[40] One way that social sanctions might operate is by making it easier for owners to collect damages from drivers since damages may often be identifiable but not legally recoverable.

REFERENCES

Ahlin, Christian, and Robert M. Townsend. 2007. "Using Repayment Data to Test across Models of Joint Liability Lending." *Economic Journal*, 117(517): F11–51.

Baker, George P., and Thomas N. Hubbard. 2003. "Make versus Buy in Trucking: Asset Ownership, Job Design, and Information." *American Economic Review*, 93(3): 551–72.

Baker, George P., and Thomas N. Hubbard. 2004. "Contractibility and Asset Ownership: On-Board Computers and Governance in U.S. Trucking." *Quarterly Journal of Economics*, 119(4): 1443–79.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul. 2005. "Social Preferences and the Response to Incentives: Evidence from Personnel Data." *Quarterly Journal of Economics*, 120(3): 917–62.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul. 2009. "Social Connections and Incentives in the Workplace: Evidence from Personnel Data." *Econometrica*, 77(4): 1047–94.

Banerjee, Abhijit V., Timothy Besley, and Timothy W. Guinnane. 1994. "Thy Neighbor's Keeper: The Design of a Credit Cooperative with Theory and a Test." *Quarterly Journal of Economics*, 109(2): 491–515.

Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan. 2009. "The Miracle of Microfinance? Evidence from a Randomized Evaluation." Unpublished.

Bertrand, Marianne, Erzo F. P. Luttmer, and Sendhil Mullainathan. 2000. "Network Effects and Welfare Cultures." *Quarterly Journal of Economics*, 115(3): 1019–55.

Besley, Timothy, and Stephen Coate. 1995. "Group Lending, Repayment Incentives and Social Collateral." *Journal of Development Economics*, 46(1): 1–18.

Borjas, George J. 2000. "Ethnic Enclaves and Assimilation." *Swedish Economic Policy Review*, 7(2): 89–122.

Duflo, Esther, and Emmanuel Saez. 2003. "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment." *Quarterly Journal of Economics*, 118(3): 815–42.

Edin, Per-Anders, Peter Fredriksson, and Olof Åslund. 2003. "Ethnic Enclaves and the Economic Success of Immigrants—Evidence from a Natural Experiment." *Quarterly Journal of Economics*, 118(1): 329–57.

Ellison, Glenn. 1994. "Cooperation in the Prisoner's Dilemma with Anonymous Random Matching." *Review of Economic Studies*, 61(3): 567–88.

Fernandez, Roberto M., and Nancy Weinberg. 1997. "Sifting and Sorting: Personal Contacts and Hiring in a Retail Bank." *American Sociological Review*, 62(6): 883–902.

Foster, Andrew D., and Mark R. Rosenzweig. 1994. "A Test for Moral Hazard in the Labor Market: Contractual Arrangements, Effort, and Health." *Review of Economics and Statistics*, 76(2): 213–27.

Gil, Ricard, and Wesley R. Hartmann. Forthcoming. "Airing Your Dirty Laundry: Vertical Integration, Reputational Capital, and Social Networks." *Journal of Law, Economics, & Organization*.

Giné, Xavier, and Dean S. Karlan. 2006. "Group Versus Individual Liability: A Field Experiment from the Philippines." Yale University Economic Growth Center Discussion Paper 940.

Giné, Xavier, and Dean S. Karlan. 2009. "Group Versus Individual Liability: Long Term Evidence From Philippine Microcredit Lending Groups." http://karlan.yale.edu/p/GroupversusIndividual-May2009.pdf.

Glaeser, Edward L., David I. Laibson, José A. Scheinkman, and Christine L. Soutter. 2000. "Measuring Trust." *Quarterly Journal of Economics*, 115(3): 811–46.

Granovetter, Mark. 1974. *Getting a Job: A Study of Contracts and Careers*. Chicago: University of Chicago Press.

Greif, Avner. 1993. "Contract Enforceability and Economic Institutions in Early Trade: the Maghribi Traders' Coalition." *American Economic Review*, 83(3): 525–48.

Holmstrom, Bengt. 1979. "Moral Hazard and Observability." *Bell Journal of Economics*, 10(1): 74–91.

Jackson, C. Kirabo, and Elias Bruegmann. 2009. "Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers." *American Economic Journal: Applied Economics*, 1(4): 85–108.

Jackson, C. Kirabo, and Henry S. Schneider. 2011. "Do Social Connections Reduce Moral Hazard? Evidence from New York City Taxi Industry: Dataset." *American Economic Journal: Applied Economics*. http://www.aeaweb.org/articles.php?doi=10.1257/app.3.3.244.

Jackson, C. Kirabo. 2010 "Match Quality, Worker Productivity, and Worker Mobility: Direct Evidence From Teachers." National Bureau of Economic Research Working Paper 15990.

Kandori, Michihiro. 1992. "Social Norms and Community Enforcement." *Review of Economic Studies*, 59(1): 63–80.

Karlan, Dean S. 2005. "Using Experimental Economics to Measure Social Capital and Predict Financial Decisions." *American Economic Review*, 95(5): 1688–99.
Karlan, Dean S. 2007. "Social Connections and Group Banking." *Economic Journal*, 117(517): F52–84.
Karlan, Dean, Markus Mobius, Tanya Rosenblat, and Adam Szeidl. 2009. "Trust and Social Collateral." *Quarterly Journal of Economics*, 124(3): 1307–61.
Lazear, Edward P. 2000. "Performance Pay and Productivity." *American Economic Review*, 90(5): 1346–61.
Luttmer, Erzo F. P. 2001. "Group Loyalty and the Taste for Redistribution." *Journal of Political Economy*, 109(3): 500–528.
Mas, Alexandre, and Enrico Moretti. 2009. "Peers at Work." *American Economic Review*, 99(1): 112–45.
Montgomery, James D. 1991. "Social Networks and Labor-Market Outcomes: Toward an Economic Analysis." *American Economic Review*, 81(5): 1407–18.
Munshi, Kaivan. 2003. "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market." *Quarterly Journal of Economics*, 118(2): 549–99.
New York City Taxi & Limousine Commission. 2005, 2007. "New York City Yellow Taxi Driver, Owner, and Taxi-Level Records." Dataset (obtained May 2005 and December 2007).
Pellizzari, Michele. 2010. "Do Friends and Relatives Really Help in Getting a Good Job?" *Industrial and Labor Relations Review*, 63(3): 494–510.
Pitt, Mark M., and Shahidur R. Khandker. 1998. "The Impact of Group-Based Credit Programs on Poor Households in Bangladesh: Does the Gender of Participants Matter?" *Journal of Political Economy*, 106(5): 958–96.
Portes, Alejandro, and Julia Sensenbrenner. 1993. "Embeddedness and Immigration: Notes on the Social Determinants of Economic Action." *American Journal of Sociology*, 98(6): 1320–50.
Schneider, Henry. 2010 "Moral Hazard in Leasing Contracts: Evidence from the New York City Taxi Industry." *Journal of Law and Economics*, 53(4): 783-805.
Sharma, Manohar, and Manfred Zeller. 1997. "Repayment Performance in Group-Based Credit Programs in Bangladesh: An Empirical Analysis." *World Development*, 25(10): 1731–42.
Stiglitz, Joseph E. 1990. "Peer Monitoring and Credit Markets." *World Bank Economic Review*, 4(3): 351–66.
Urbanomics. 2004. *Taxi & Limousine Commission Environmental Impact Statement Socioeconomic Conditions Report*. Internal document prepared for the Taxi & Limousine Commission.
Wydick, W. Bruce. 1999. "Credit Access, Human Capital and Class Structure Mobility." *Journal of Development Studies*, 35(6): 131–52.

This article has been cited by:

1. R. Gil, J. Marion. 2012. Self-Enforcing Agreements and Relational Contracting: Evidence from California Highway Procurement. *Journal of Law, Economics, and Organization* . [CrossRef]