

Balafoutas, Loukas; Beck, Adrian; Kerschbamer, Rudolf; Sutter, Matthias

Working Paper

What drives taxi drivers? A field experiment on fraud in a market for credence goods

CESifo Working Paper, No. 3461

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Balafoutas, Loukas; Beck, Adrian; Kerschbamer, Rudolf; Sutter, Matthias (2011) : What drives taxi drivers? A field experiment on fraud in a market for credence goods, CESifo Working Paper, No. 3461, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/46536>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

What Drives Taxi Drivers? A Field Experiment on Fraud in a Market for Credence Goods

Loukas Balafoutas
Adrian Beck
Rudolf Kerschbamer
Matthias Sutter

CESIFO WORKING PAPER NO. 3461
CATEGORY 13: BEHAVIOURAL ECONOMICS
MAY 2011

An electronic version of the paper may be downloaded

- *from the SSRN website:* www.SSRN.com
- *from the RePEc website:* www.RePEc.org
- *from the CESifo website:* www.CESifo-group.org/wp

What Drives Taxi Drivers? A Field Experiment on Fraud in a Market for Credence Goods

Abstract

Credence goods are characterized by informational asymmetries between sellers and consumers that invite fraudulent behavior by sellers. This paper presents the results of a natural field experiment on taxi rides in Athens, Greece, set up to measure different types of fraud and to examine the influence of passengers' presumed information and income on the extent of fraud. Results reveal that taxi drivers cheat passengers in systematic ways: Passengers with inferior information about optimal routes are taken on longer detours while asymmetric information on the local tariff system leads to manipulated bills. Higher income seems to lead to more fraud.

JEL-Code: C930, D820.

Keywords: credence goods, expert services, natural field experiment, taxi rides, fraud, asymmetric information.

Loukas Balafoutas
University of Innsbruck
Austria / Innsbruck
loukas.balafoutas@uibk.ac.at

Adrian Beck
University of Innsbruck
Austria / Innsbruck
adrian.beck@uibk.ac.at

Rudolf Kerschbamer
University of Innsbruck
Austria / Innsbruck
rudolf.kerschbamer@uibk.ac.at

Matthias Sutter
Department of Public Finance
University of Innsbruck
Universitaetsstrasse 15
Austria – 6020 Innsbruck
matthias.sutter@uibk.ac.at

This version: 28th April 2011

We thank Tim Barmby, Colin Camerer, Gary Charness, Nick Feltovich, Uri Gneezy, Martin Kocher, Charles Noussair, and Jan Potters, as well as audiences at APESA 2011 in Kuala Lumpur, IMEBE 2011 in Barcelona, RES 2011 in London and the universities of Aberdeen, Athens, Edinburgh, Innsbruck, Lyon, Tilburg, and York for helpful comments. Special thanks are due to Konstantinos Konstantakis for assisting in the experiment. Financial support from the Austrian Science Fund (FWF grant P20796) is gratefully acknowledged.

I. INTRODUCTION

Many goods and services, such as car or computer repairs, medical treatments or taxi rides in an unknown city, share the characteristic of an informational asymmetry between the seller and the consumer. Typically, the seller knows more than the consumer about the quality that yields the highest surplus from trade, and moreover the consumer is frequently not even *ex post* able to observe the quality of the good or service he received. . These features have led Darby and Karni (1973) to create the term “credence goods”, because consumers are not able to monitor quality.

Such informational asymmetries open the door to a whole array of different types of fraudulent behavior on the sellers’ side (Dulleck and Kerschbamer, 2006), including *overtreatment* (providing a higher quality than the surplus maximizing one), *undertreatment* (choosing a quality that is insufficient to satisfy the consumer’s needs), and *overcharging* (charging for a higher quality than has been provided). Fraud in markets for credence goods has potentially large efficiency costs for an economy. For instance, the U.S. Department of Transportation has estimated that more than half of the total expenses for car repairs are for unnecessary repairs, which can be taken as a rough estimate of the efficiency costs of overtreatment (Wolinsky, 1993). Referring to the health care sector, a Swiss study has found that the average person’s probability of receiving one of seven major surgical interventions is one third above that of a physician or a member of a physician’s family, indicating that a consumer’s (presumed) education or information level affects the quality of the received treatment and therewith the likelihood of under- and overtreatment (Domenighetti et al., 1993). A Japanese study has revealed that doctors’ prescriptions respond to differences in mark-up across different drugs, i.e., to monetary incentives that are unrelated to warranted medication (Iizuka, 2007).

Yet, the complexity of services such as medical treatments or car repair services remains an obstacle to unambiguously measuring the extent of fraud. In fact, “mistreatment” in such markets could be driven by causes unrelated to fraud. For instance, an overly cautious car mechanic (physician) might be prone to conducting repairs (doing check-ups) prematurely in the best intention to prolong the lifetime of a car or patient, respectively. Moreover, it is difficult to detect and quantify the amount of overcharging through sellers charging for services that have not been provided.¹

¹ In the case of car repairs, overcharging can occur through inflating the time spent on the repair or charging for goods and services that have not been provided at all (Dulleck and Kerschbamer, 2006). While in the health

From a policy perspective, a better understanding of the extent of fraud and its determinants in credence goods markets is important in order to design appropriate incentive structures or regulatory frameworks. In this paper, we present the results of a natural field experiment designed to identify the extent *and* type of fraud and their determinants in a market for a credence good. More precisely, our field experiment focuses on the market for taxi rides. In this market, confounding factors present in other credence goods markets – such as issues of misdiagnosis or incompetence of the expert seller, overly cautious behavior, or the inexistence of services that solve a consumer’s problem – are practically non-existent. In addition, undertreatment by failing to reach the requested destination is hardly an issue, and overtreatment is easy to identify for an informed external observer since unnecessary detours are unlikely to add any value for the customer. Given those features of the market for taxi rides, the present paper contributes to the literature on fraudulent behavior on markets for credence goods in two substantial ways.

First, running a field experiment in a market for taxi rides allows us to identify the extent of fraud *and* to disentangle it into the two dimensions *overtreatment* and *overcharging*. We do so by letting three experimenters take taxi rides in the capital city of Greece, Athens, a city with approximately 4 million inhabitants and 14,000 taxis. In total, the experimenters took 174 rides, with a total driving distance of more than 2,200 km and an overall duration of 63 hours of taxi driving. For each single ride, a portable GPS satellite logger allowed us to precisely record the chosen route and the taxi’s exact position and speed at each second of the ride. With these data, we can quantify overtreatment in the form of detours. The GPS data are also the key to measuring overcharging, because they allow calculating the correct fare for a given distance. The difference between the total fare charged by the taxi driver and the correct fare then measures the amount of overcharging.

Second, and in our view even more important, we study how the extent and the type of fraud in the market for taxi rides depend on key characteristics of a passenger. In our experimental treatments, we manipulate the taxi driver’s perception about the passenger’s (i) information about the city and the optimal route, (ii) information about the tariff system, and (iii) income.

care sector overcharging can occur in analogous ways, the incentives for overcharging are exacerbated there since overcharging is of little immediate concern to consumers if obligatory social security systems bear the costs of medical treatment. Overcharging is a problem for insurers, however, as they have no easy means of controlling what kind of service has actually been provided to patients.

We let the three experimenters always ask for the same service at almost the same time by departing one by one from the same starting point to the same destination. In each triple, two of the experimenters spoke in Greek to the taxi driver, while the third one spoke in English. An experimenter in the role of an English speaking passenger and one experimenter in the role of a Greek speaking passenger stated the destination and asked whether the driver knew the address, adding as an explanation for asking that they were not familiar with the city. The other Greek speaking passenger, however, only stated the destination, but did not say or ask anything else. The purpose of this treatment variation was to manipulate the *driver's perception of the passenger's familiarity with the city*. We hypothesized that passengers revealing that they were unfamiliar with the city would be more prone to being taken on detours.

Varying the passenger's language (Greek versus English) was meant to manipulate a *driver's perception of the passenger's familiarity with the details of the local taxi tariff system*. Since taxi tariffs are subject to the same regulation all over Greece, speaking Greek was meant to convey to the driver that the passenger is likely to know the general rules for charging, such as when the daytime-tariff applies, or what kind of surcharges are allowed for which services. Since an English speaking passenger is arguably less likely to be perceived as familiar with the details of the Greek taxi tariff system, we hypothesized that taxi drivers try to exploit their informational advantage in the charging dimension more extensively with English speaking than with Greek speaking passengers.²

Finally, we tried to manipulate a *driver's perception of the passenger's income* by varying the passenger's clothes and the requested destination. Passengers in the *high income* role always wore a suit and carried a briefcase, while passengers in the *low income* role were dressed in casual clothes and carried a backpack. If the respective destination was a hotel, it was either an expensive top-end hotel for high-income passengers or a cheap hostel for low-income passengers. Assuming that taxi drivers have convex distributional preferences, our hypothesis was that high-income passengers were more likely to be cheated upon by taxi drivers.

² It is a long established pattern in linguistics that people tend to infer (correctly or incorrectly) private information from speech. In a classic field experiment, Kingsbury (1968) asked randomly selected pedestrians on a Boston street for directions to a department store. He asked his question in either the local Boston dialect, or employing a dialect spoken in rural Missouri. The results showed that in the latter case the directions given were significantly longer and more detailed than in the former case. Apparently, on the basis of his foreign dialect alone, people assumed that his level of local expertise was low, and reacted on that in their responses.

Our field experiment provides a number of advantages compared to other approaches to examine expert sellers' behavior on credence goods markets.³ Compared to non-experimental field studies like those discussed on the opening page, it allows for a systematic variation of treatment variables, while still investigating the behavior of real expert sellers in their natural working environment. Compared to laboratory experiments, external validity is practically of no concern in a natural field experiment.⁴

We are aware of only one previous field experiment in a credence goods market. Schneider's (2009) study on fraudulent behavior in the car repair sector focuses on reputation concerns, finding no influence of such considerations on the service provided, however. While Schneider (2009) reports that overtreatment and undertreatment are pervasive, he acknowledges that he "cannot rule out incompetence as a factor contributing to under- and overtreatment" (p. 27). Moreover, unnecessary repairs might have been caused by overly cautious mechanics who replaced parts to avoid possible malfunctioning in the near future. As mentioned earlier, such potential confounds are minimized in the market for taxi rides.⁵

We now turn to a short preview of our main findings. The average length of a ride was 12.9 kilometers (km), of which on average 1 km (8%) was an unnecessary detour. Those passengers who revealed their unfamiliarity with the city were, on average, taken on detours of double length compared to passengers who did not. Overall, in 11% of cases passengers were overcharged through the application of incorrect tariffs. English speaking passengers, who probably conveyed the impression of not being familiar with the details of the tariff system, were overcharged in 19% of the cases, while this happened in only 7% of cases to Greek speaking passengers. Concerning the income dimension, we find that passengers in the role of having high income are taken on longer detours and experience slightly more overcharging.

The rest of the paper is organized as follows. In Section II we provide some background information on the taxi market in Athens, thus setting the stage for our field experiment. Section III introduces the experimental design and derives our hypotheses.

³ See List (2006) and List and Reiley (2008) for surveys on the advantages of field experiments.

⁴ For instance, one of the findings of the laboratory experiments by Kerschbamer et al. (2009) and Dulleck et al. (2011) is that sellers' distributional preferences have a large impact on their provision and charging decisions. It is far from clear, however, how this result translates to real world credence goods markets. By manipulating drivers' perception of the income of passengers in our natural field experiment we can address this question.

⁵ There is an interesting literature on the labor supply of (New York City) taxi drivers (see, e.g., Camerer et al., 1997; Farber, 2005, 2008; Crawford and Meng, 2011). However, this literature remains silent on the question of whether and to which extent taxi drivers exploit their informational advantage over customers in the provision of this credence good. Hence, we are not going into the details of these studies.

Section IV presents the results on the determinants of different types of fraudulent behavior in this market. Section V discusses our findings along with different explanations for them. Section VI concludes the paper.

II. THE MARKET FOR TAXI RIDES IN ATHENS

The market for taxi rides is regulated nationwide in Greece. The regulation concerns both market entry and the tariff system.⁶ Market entry is regulated through taxi licenses that are issued by a governmental authority and valid for a lifetime.⁷ At the time of running the field experiment (in July 2010), approximately 14,000 taxi licenses were valid in Athens. This means that there were roughly 350 taxis per 100,000 inhabitants. This quota is noticeably larger than in some other cities. For instance, in London the comparable number is 280 taxis, in Berlin 210 taxis, and across the U.S. the overall average is around 110 taxis per 100,000 inhabitants.⁸ This comparison shows that the supply of taxi rides is relatively large in Athens.

The tariff system is regulated such that there is a fixed fee for entering a taxi plus a variable fee, dependent on the time of the day, the distance traveled and the duration of a ride. This corresponds to a time-varying two-dimensional two-part tariff, a tariff type in place also in many other major cities, for instance in New York City. All over Greece, the tariff looks as follows (effective July 2010): On top of a fixed fee of €1.16, the distance-dependent tariff yields €0.66 per kilometer during daytime (i.e., from 5 a.m. until midnight) and €1.16 per kilometer during nighttime. By contrast, the duration-dependent tariff is invariable and yields €0.1775 per minute. The algorithm for charging is standardized nationwide in all taximeters and it switches automatically to the counting method (distance-dependent vs. duration-dependent) that is more profitable for the driver. That is, if the vehicle's speed during daytime is above 16 kilometers per hour (km/h), only distance is charged, while below 16 km/h only duration is charged. Under reasonable assumptions, it is then straightforward to show that it is more profitable for a taxi driver to take a passenger on a detour than to choose a shorter route

⁶ Many major cities also regulate both market entry and the tariff system. Examples include New York City, Paris, Brussels, Helsinki, and Quebec, to name just a few (see The New York City Taxicab Fact Book, 2006, Bekken, 2003, OECD, 2007, http://www.iedm.org/uploaded/pdf/aout2010_en.pdf).

⁷ Lifelong licenses are also issued in Boston, Chicago, New York City, or Melbourne, for example (see Schaller, 2005), while they have to be renewed in London, for instance (see <http://www.tfl.gov.uk/businessandpartners/taxisandprivatehire/1346.aspx#section-2>).

⁸ Source: Own calculations, based on numbers from The New York City Taxicab Fact Book (2006), "Transportation for London" (<http://www.tfl.gov.uk/businessandpartners/taxisandprivatehire/1380.aspx>), "Taxi Innung Berlin" (<http://www.taxiinnung.org/Taxi-Bestellen.24.0.html>), and Schaller (2005).

on which he is likely to get stuck in a traffic jam.⁹ As far as *overcharging* is concerned, the incentives are straightforward, since overcharging increases the driver's revenue without affecting the cost of service. It is important to stress that the marginal incentives of taxi drivers to engage in fraud of any type do not depend on whether the driver owns or leases the vehicle, because leasing a taxi comes at a fixed cost (of roughly €35 per shift; see www.satataxi.gr) and no part of the revenue collected during the shift is shared with the owner. Therefore, taxi drivers are always residual claimants, meaning that the possible profits from fraud in the two dimensions overtreatment and overcharging are reaped by themselves.¹⁰

III. THE FIELD EXPERIMENT

A. Method and Procedure

The experiment involved three experimenters switching between different passenger roles. In order to minimize the potential for confounding effects of a passenger's age or gender on driver behavior, all three experimenters were male and in the late twenties. For each route, the three experimenters always took a ride from the same starting point to the same destination. They approached the taxi stand one by one in intervals of one to two minutes, so that taxi drivers never saw them together. The short intervals were chosen in order to control for a host of unforeseeable factors, such as variations in traffic, road works, or accidents, that may influence the optimal route. In the following, we denote the three rides that were taken at practically the same time as a *triple*.

Each experimenter always took a seat in the back of the taxi and he was equipped with a GPS satellite logger. This small device (see Picture A.1 in the Appendix) is easy to hide and allowed the experimenters to record the exact route driven, the total distance travelled, the exact duration of the ride, and the location and speed of the taxi at each point in time. In

⁹ Assuming an average speed of 40 km/h, a fuel consumption of 8.0 liters per 100 km while driving and 1 liter per hour while waiting, and a gas price of 1.3 € per liter, a minute spent on a detour during daytime yields €0.37 [= (0.67 km/min x 0.66 €/km) - (0.67 km/min x 0.08 l/km x 1.3 €/l)], which is considerably more than the estimated earnings of €0.16 [= (0.1775 €/min - 160 l/min x 1.3 €/l)] for a minute spent in a traffic jam. The incentives for overtreatment are even stronger during nighttime due to the more attractive distance-dependent night tariff, but in this study we collected observations only during daytime to keep the drivers' incentives constant.

¹⁰ This is often not the case in other credence goods markets. For example, car mechanics or doctors are frequently employed for a fixed wage, which weakens their direct financial incentives for overtreatment and overcharging, thus making it much more difficult to observe and measure those types of fraud.

addition to the GPS data, the experimenters collected data on the total fare, and the sex and approximate age of the driver.

Our treatment variations were implemented as follows. In each triple of rides, the three experimenters were in three different “information roles”. We refer to them in the following as *local*, *non-local native*, and *foreigner*, respectively. In all three roles, an experimenter instructed the driver upon entering the taxi to take him to a particular destination. Passengers in the roles of locals and non-local natives did this in Greek, while passengers in the role of foreigners gave the instructions in English. Passengers in the role of non-local natives and of foreigners then asked the driver whether he knew the destination, adding as an explanation for asking that they were not familiar with the city. The question whether the driver knew the destination (plus the added explanation) is the only difference between locals and non-local natives, since both types of passengers talked in Greek. The language is the only difference between non-local natives and foreigners, both of whom had the same text when entering the taxi, one speaking in Greek, the other one speaking in English.

In addition to an information role, each passenger also had an “income role”. We manipulated a taxi driver’s perception of the passenger’s income by varying the passenger’s clothes and (whenever possible) additionally the destination hotel. Passengers intended to be perceived as having high income were dressed in a suit and carried a briefcase, while passengers in the role of having low income were dressed casually and carried a backpack. For those routes where the destination was a hotel, a high-income passenger would drive to a top-end hotel, while a low-income passenger would have a low-end hostel as his destination.¹¹ Table 1 summarizes our treatments and the number of observations per treatment.

[Table 1 about here]

In order to ensure randomization, we collected observations during two weeks, covering every day of the week and every time of day between 8 a.m. and midnight. For the same reason (and in order to have a representative sample) the observations were not collected on a single route, but on 15 different ones, covering large parts of Athens and including rich and poor neighborhoods, as well as typical tourist spots.¹² Within each triple of rides, the order with which passengers in the different roles approached the taxi stand and

¹¹ The two addresses were very close to each other in the same street, meaning that the distance was practically identical.

¹² Tables A.1 and A.2 in the Appendix list all destinations and routes. Picture A.2 illustrates all routes in a map.

took a taxi was randomized. On average, the three experimenters entered the taxis in intervals of 83 seconds one after the other.

B. Hypotheses

Our first hypothesis concerns the influence of the informational asymmetry regarding the optimal route to the destination on the extent of overtreatment by taking detours. Given that both non-local native and foreign passengers expressed to be unfamiliar with the city, we expected them to be more prone to be taken on detours than local passengers, because the latter did not reveal any kind of unfamiliarity with the city. We did not expect to see differences between non-local natives and foreigners since both were arguably perceived as equally poorly informed. We summarize as our first hypothesis:

H1 (Information on the City): *Non-local native passengers and foreign passengers are more prone to overtreatment than local passengers. The extent of overtreatment does not differ between non-local native passengers and foreign passengers.*

Our second hypothesis refers to the influence of the informational asymmetry regarding the details of the tariff system on the amount of overcharging. Passengers in the role of locals and non-local natives were likely to be perceived as reasonably well, and equally well, informed about the details of the tariff system, because the tariff system is identical across the whole nation. Foreign passengers were less likely to have such information, for instance, when the night-time tariff applies, whether or not surcharges are due for a given service, or that the driver is required by law to issue a proper receipt.¹³ This informational asymmetry makes foreign passengers more susceptible to overcharging, as stated in our second hypothesis.

H2 (Information on Tariffs): *Foreign passengers are more prone to overcharging than local passengers and non-local native passengers. The extent of overcharging does not differ between local passengers and non-local native passengers.*

Our third hypothesis is motivated by the large evidence from laboratory experiments that distributional preferences are behaviorally relevant in many important market and non-

¹³ See Picture A.3 in the Appendix for a sample receipt.

market transactions (see Cooper and Kagel, 2011, for a recent review). Convex distributional preferences imply that a decision maker's benevolence towards another individual increases (or that malevolence decreases) as the income of the other individual decreases along an indifference curve (see Cox et al., 2008, for details). This is independent of the archetype to which a decision maker's distributional preferences belong, i.e., whether the decision maker is altruistic, inequality averse, spiteful, maximin, or whatsoever. Since the overwhelming majority of decision makers who are not exclusively interested in the maximization of their own material income has convex distributional preferences (for experimental evidence see, e.g., Fehr and Schmidt, 1999, Bolton and Ockenfels, 2000, Andreoni and Miller, 2002, or Charness and Rabin, 2002), we expected taxi drivers to overtreat or overcharge low-income passengers less than high-income ones. This is recorded as our third hypothesis:

H3 (Income): High-income passengers are more prone to overtreatment and overcharging than low-income passengers.

IV. RESULTS

In total, the three experimenters took 174 taxi rides, which added up to 2,236 km of travelling through Athens and 63 hours of driving. The total cost for all rides was €2,112. On average, a ride was 12.9 km long, lasted for 22 minutes, and cost €12.14. All except five taxi drivers in our sample were male (97%). Each single ride ended at the requested destination, meaning that undertreatment by failing to reach the destination was not an issue.¹⁴

A. Overtreatment

In order to get a measure of overtreatment, we calculate an *Overtreatment Index* by taking for each triple of rides the shortest trip and normalizing the other two trips by the shortest one. Table 2 presents the results, showing an index of 1.03 for locals, and one of 1.08 both for non-local natives and foreigners. The difference between passengers in the role of locals and each of the other two passenger roles is statistically significant ($p < 0.05$; two-sided

¹⁴ Note that we purposefully chose destinations that are easy to find, and hence taxi drivers always knew where the destination was. In case the driver asked which route to take, the choice was explicitly left to the driver.

Wilcoxon signed ranks tests¹⁵). There is no significant difference between non-local natives and foreigners, however. These results support our hypothesis H1, because they show that passengers revealing not to be familiar with the city are taken on significantly longer detours.

[Table 2 and Figure 1 about here]

Figure 1 plots the cumulative distribution function (cdf) of the Overtreatment Index for each of the three information roles. It confirms that local passengers experience significantly less overtreatment than both non-local natives ($p < 0.05$, two-sided Kolmogorov-Smirnov test) and foreign passengers ($p < 0.01$), but there is no significant difference between non-local natives and foreigners ($p > 0.4$).

In the Appendix (Table A.3 and Figure A.1) we present – as an additional control and robustness check – a differently calculated measure for overtreatment that yields exactly the same insights. For this purpose, we first determined for each ride the shortest possible route. We did this by driving all the routes ourselves after the end of the experiment, following the directions of the GPS navigation system “TomTom”. We then calculated an *Alternative Overtreatment Index* that divides the actually driven distance in a particular ride (derived from the GPS data) by the shortest possible distance. This index is 1.08 across all rides, implying that a ride’s average length of 12.9 km included a detour of about 1 km. Detours accounted for at least 5% of the shortest possible routes in 45% of rides. The Alternative Overtreatment Index is 1.05 for local passengers, but 1.10 for both, non-local native and foreign passengers ($p < 0.05$ when testing locals against each of the other two roles; two-sided Wilcoxon signed ranks tests). Hence, the results on hypothesis H1 are robust to this alternative specification.

Of course, a longer and more costly route could, in principle, be driven by the desire of a driver to save on the passenger’s time by taking a quicker, albeit longer, route. This is not what we find in the data. In fact, the average duration of a ride was shortest for locals (21 minutes and 17 seconds), intermediate for non-local natives (21:38 min), and longest for foreigners (22:11 min). Figure 2 presents the cumulative distribution function of a *Duration Index*, calculated as follows. We normalize each ride in a given triple by the quickest ride in that triple. The horizontal axis in Figure 2 therefore measures the relative duration of a ride in

¹⁵ For these tests, we match the observation for the local passenger with the non-local native passenger (respectively foreign passenger) in a triple of rides and apply the Wilcoxon signed ranks test to pairs of matched observations in all triples.

comparison to the quickest ride. While the figure reveals a tendency for longer rides with foreigners, the differences in the distributions among the three passenger roles are not significant according to a Kolmogorov-Smirnov test.¹⁶ Hence, to say the least, the differences in the length of detours shown in Figure 1 are certainly not compensated by shorter travel times of those passengers that are being taken on longer detours.

[Figure 2, Figure 3 and Figure 4 about here]

Turning to the influence of the income role, Table 2 shows that high-income passengers have an average Overtreatment Index of 1.09, and low-income passengers an index of 1.05 ($p < 0.05$; two-sided Mann-Whitney U-test). Figure 3 presents the cumulative distribution function, and it confirms the smaller Overtreatment Index for low income passengers ($p < 0.05$; two-sided Kolmogorov-Smirnov test). Figure 4 shows the cumulative distribution function of the Duration Index for high-income and low-income passengers, revealing that the index is clearly higher for high-income passengers ($p < 0.05$; two-sided Kolmogorov-Smirnov test). This implies that high-income passengers are not only taken on longer detours, but it also costs them more time. In sum, the aggregate data on the influence of a passenger's presumed income support hypothesis H3.

B. Overcharging

Overcharging occurs when a passenger pays more than he should *for a given distance*. Since the actually driven distance could be reconstructed from the GPS data, thus allowing to control for overtreatment, overcharging was identified in our data if one of the following events occurred: (i) the driver switched to the more profitable night tariff even though the ride was during daytime; (ii) the driver did not switch on the taximeter at all, but at the end of the ride demanded a higher price than justified by the distance traveled and the duration of the ride; or (iii) the driver demanded an amount higher than the one shown on the taximeter, with the justification of bogus surcharges.

In total, we observed overcharging in 19 out of 174 rides (11%). The sources of overcharging were the following. In only one case the taximeter was not switched on, six

¹⁶ Testing with a Jonckheere-Terpstra-Test whether the rides in a triple are ordered, such that they are quickest for local passengers, intermediate for non-local natives, and slowest for foreign passengers, yields $p = 0.078$, however.

cases were due to the unjustified usage of the night tariff, and bogus surcharges accounted for twelve cases. Despite the relatively low overall frequency of overcharging, there are clear differences across passengers' information roles, as can be seen in panel [A] of Table 3. While overcharging occurred in only 5% (9%) of rides with local (non-local native) passengers, it happened in 19% of cases to foreigners ($p < 0.05$ for locals vs. foreigners; $p = 0.1$ for non-local natives vs. foreigners; χ^2 -tests).

[Table 3 about here]

Panel [B] of Table 3 presents the average amount of overcharging by a passenger's information and income role, conditional on overcharging having taken place. The average amount of overcharging in the 19 relevant cases was €4.30, which corresponds to 35% of the average fare. Concerning treatment differences, we note that the amount of overcharging is highest with foreign passengers. They face significantly higher amounts of overcharging when we test against the pooled data of Greek-speaking passengers, i.e., locals and non-local natives ($p < 0.05$; two-sided Mann-Whitney U-test; $N = 19$). Controlling for a passenger's information role, overcharging is also higher for high-income passengers than low-income passengers.

Overall, the results on overcharging support our hypothesis H2 that predicted (i) more overcharging for passengers in the role of foreigners, and (ii) no differences between local and non-local native passengers. The data are also largely consistent with hypothesis H3, because high-income passengers pay, on average, more as a consequence of overcharging than low-income passengers. While the difference is not significant with non-parametric tests (given the low number of observations), the regressions below will show a weakly significant effect of income on overcharging.

C. Total Fare

Now we take a look at the total fare in the different treatments as an indicator of the overall amount of fraud. Table 4 presents a *Fare Index* that is calculated as the ratio of a passenger's fare over the minimum fare among the three rides in a triple. In line with hypotheses H1 and H2, we see that non-local natives paid higher fares than locals (Fare Index of 1.09 vs. 1.04; $p < 0.01$; two-sided Wilcoxon signed ranks test), and that foreigners paid higher fares than non-local natives (Fare Index of 1.19 vs. 1.09; $p = 0.05$). The former result is largely driven by differences in overtreatment, while the latter is mainly due to differences in

overcharging. On average, the Fare Index is higher for high-income passengers than low-income passengers, but non-parametric tests fail conventional significance levels.

[Table 4 about here]

D. *Econometric Analysis*

Table 5 presents five different Tobit regressions. In the top row we indicate the dependent variables. As independent variables we use, first, a dummy for having expressed unfamiliarity with the city ($\neg local$). This dummy captures passengers both in the role of non-local natives and foreigners, and thus their presumed informational disadvantage concerning the optimal route to a destination. Second, we include a dummy for passengers speaking in English (*foreign*). This dummy is intended to reflect the effects of being less likely familiar with the tariff system than Greek-speaking passengers. Third, we add a dummy for high-income passengers (*high income*). In specifications (2) and (5) we also include interaction terms of *high income* and $\neg local$, respectively *foreign*. In addition to these variables we insert further controls, such as the driver's gender and approximate age and the *time of day* (measured in one-hour intervals from 8 a.m. to midnight).¹⁷ To account for the fact that our observations are not independent of each other within each triple of rides, we cluster standard errors by set of three simultaneous rides.

[Table 5 about here]

Specification (1) presents a Tobit regression on the Overtreatment Index. The group of local passengers constitutes the benchmark group. We see that $\neg local$ is significant and adds an estimated 9.6% of detours. This means that revealing unfamiliarity with the city increases overtreatment considerably. Being perceived as a foreign passenger does not further add to overtreatment (in comparison to a non-local native), which is in line with our hypothesis H1. High-income passengers are taken on significantly longer detours than low-income passengers, with an estimated difference of 6.1%. This is in line with hypothesis H3. Specification (2) of Table 5 shows that including interaction terms of *high income* and $\neg local$,

¹⁷ We also tried specifications including route and day fixed effects. We dropped the fixed effects based on F tests of joint significance of the day and route dummies ($p > 0.6$ for day fixed effects, $p > 0.8$ for route fixed effects).

respectively *foreign*, does not change the significance and direction of the main effects, while the interaction terms themselves are insignificant. However, the main effect of *high income* is restricted to local passengers, because the interaction terms offset the main effect of *high income* for non-local native and foreign passengers.¹⁸

Specification (3) presents the results of a Tobit regression on the amount of overcharging (in €). While conveying the impression of unfamiliarity with the city does not have a significant impact in itself (see " $\neg local$ "), being in the role of a foreigner increases the amount of overcharging significantly by an estimated €5.15 (see "*foreign*"). This matches our hypothesis H2, according to which passengers who are more likely to be perceived as having an informational disadvantage about the general tariff system in Greece face more overcharging. Income plays a weakly significant role for overcharging, thus providing some support for our hypothesis H3.¹⁹

Finally, specifications (4) and (5) report Tobit regressions on the Fare Index, once with interaction terms, once without. Here we find that both, $\neg local$ and *foreign*, are significantly positive. This shows that both dimensions of informational asymmetries (on the optimal route and on the tariff system) contribute to a higher fare. The effect of *high income* is weakly significantly positive in specification (5), but restricted to local passengers, because the interaction terms offset the main effect of *high income* for non-local native and foreign passengers.²⁰

E. Summary of Main Results

We conclude the results section by formulating three results, organized along the lines of our hypotheses:

Result 1 (Information on the City): Consistent with hypothesis H1, we find that (i) non-local native passengers and foreign passengers are more prone to overtreatment than local passengers, and that (ii) there is no significant difference between non-local natives and foreigners in the extent of overtreatment.

¹⁸ Wald tests yield $p = 0.25$ on the restriction $high\ income + \neg local \times high\ income = 0$; $p = 0.94$ on the restriction $high\ income + \neg local \times high\ income + foreign \times high\ income = 0$.

¹⁹ Due to the small number of observations on overcharging, the variable *driver female* had to be dropped here and another specification that adds interaction terms to the overcharging Tobit regression was infeasible.

²⁰ Wald tests yield $p = 0.17$ on the restriction $high\ income + \neg local \times high\ income = 0$; $p = 0.15$ on the restriction $high\ income + \neg local \times high\ income + foreign \times high\ income = 0$.

Result 2 (Information on Tariffs): Consistent with hypothesis H2, we find that (i) foreign passengers are more prone to overcharging than local passengers and non-local native passengers, and that (ii) there is no significant difference between locals and non-local natives in the frequency and extent of overcharging.

Result 3 (Income): There is only weak support for hypothesis H3. While high-income passengers are, on average, more prone to overtreatment and overcharging than low-income passengers, the main effects are only weakly significant, and they apply only to local passengers.

V. DISCUSSION

A. Explaining Treatment Differences Across Information Roles

Differences in Detection Probabilities. Our hypotheses H1 and H2 can be derived by assuming that taxi drivers suffer a cost if they are detected as cheaters. Given such a cost, differences in passengers' perceived information and the associated differences in detection probabilities translate into differences in the incentives for fraud. Hypothesis H1 relies on the assumption that the probability of overtreatment being detected is lower for passengers in the role of non-local natives or foreigners than for locals, while there is not much difference in this probability between non-local natives and foreigners. Hypothesis H2, by contrast, supposes passengers in the role of foreigners are perceived as less likely to detect overcharging than passengers in the role of locals and non-local natives, while there is not much difference in the perceived probability to detect overcharging between locals and non-local natives.

To us, it seems plausible that a cost of detection exists in material or non-material form. Detection may have a material cost through penalties for fraudulent behavior, either as monetary fines or as the loss of a driver's taxi license, thus expelling the driver from the business. It is important to note that there is an asymmetry between overtreatment and overcharging with respect to potential material costs. A driver accused of overtreatment can always argue that he took a detour in the best interest of the passenger, for instance, to evade a traffic jam. There is no such excuse for overcharging. Thus, potential material costs are less likely to deter overtreatment than overcharging. Indeed, overtreatment was observed much more frequently than overcharging. While 45% of passengers were taken on detours that accounted for at least 5% of the shortest route, only 11% of passengers were overcharged.

Besides the potential material costs, there may be non-material utility losses from being detected as a cheater. These may arise from unpleasant discussions with angry passengers about the route taken or the tariff applied. Moreover, a driver may feel ashamed or guilty if a passenger finds out that he was overtreated or overcharged. The theory of guilt aversion (Charness and Dufwenberg, 2006), for instance, implies that guilt averse subjects behave as if they have a cost of cheating on others.

Differences in Reporting Probabilities. There is an alternative explanation for parts of our data patterns that relies on differences in reporting rather than differences in detection probabilities. Assume that the costs of reporting fraud to the authorities are smallest for local passengers, intermediate for non-local natives and highest for foreigners. Such an ordering arises, for instance, when locals know where to report, while non-local natives have to find out, and foreigners additionally face a severe language problem in communicating with the authorities. Under these assumptions one would expect the lowest level of fraud for local passengers, an intermediate level for non-local natives, and the highest level for foreigners. In fact, this matches our observations concerning the Fare Index in Table 4. However, this alternative explanation is hard to reconcile with the disentangled data on overtreatment and overcharging. Specifically, there is a treatment difference in *overtreatment* between locals, on the one hand, and non-local natives and foreigners on the other hand. There is no treatment difference, though, in overtreatment between non-local natives and foreigners. At the same time, there is a treatment difference in *overcharging* between the set of locals and non-local natives and the set of foreigners, yet no treatment difference in overcharging between locals and non-local natives. This pattern is hardly consistent with an explanation relying on differences in reporting probabilities, because differences in those probabilities are unlikely to depend on the fraud dimension. Note that our line of reasoning in this paragraph – and the two following ones – depends on our ability to disentangle overtreatment from overcharging. Hence, the GPS data are also instrumental to discriminate between different explanations for our findings.

Reputation Concerns. An alternative explanation for the effects of a passenger's information role on the extent of fraud is related to reputation concerns of taxi drivers. The probability of repeated business might be perceived as highest for passengers in the role of locals, intermediate for non-local native passengers, and lowest for foreign passengers. If this were the case, then a driver's incentive to give up short-term gains from fraud in expectation of larger future benefits from repeated business would vary with a passenger's information role such that, on average, local passengers would face the lowest levels of fraud, non-local

native passengers an intermediate amount, and foreigners the highest amount. In fact, this is what we find for the Fare Index discussed in Table 4, but this explanation does not fit the observed patterns of overtreatment and overcharging. This is so because differences in the perceived probabilities of repeated business across passengers should not result in different behavioral implications depending on the type of fraud. Finally, it is also worth mentioning that not a single driver ever offered a business card, mobile phone number, or the like in an attempt to secure another ride with this passenger in the future, making it rather unlikely that reputational concerns are an important factor for behavior in this market at all.

In-Group Favoritism. An alternative explanation for the differences in the amount of fraud between locals, non-local natives and foreigners is in-group favoritism by taxi drivers towards local passengers and (possibly to a somewhat lesser extent) towards non-local natives.²¹ We consider this explanation as implausible. Specifically, the finding that local passengers are taken on significantly shorter detours than non-local natives, while the frequency of overcharging does not differ between both groups, combined with the finding that non-local natives are less often overcharged (but not less overtreated) than foreigners, is difficult to bring in line with a plausible in-group favoritism hypothesis that does not rely on different in-group definitions in the different domains of fraud in the market for taxi rides.

B. Explaining Treatment Differences Between Income Roles

Distributional Preferences. Our hypothesis H3 is based on the argument that decision makers with convex distributional preferences tend to behave more benevolently towards low-income than towards high-income passengers. Thus, if drivers have convex distributional preferences and if they perceive passengers in the role of high-income customers indeed as richer than low-income passengers, then they should cheat less on the low-income passengers. Result 3 provides some support for this interpretation because we have observed, on average, more overtreatment and more overcharging with high-income passengers. Therefore, distributional preferences seem to have an influence on the provision behavior of expert sellers of a credence good not only in the laboratory (Dulleck et al., 2011), but also somewhat in the field. Result 3 has revealed that the effect is not significant across all information roles, however,

²¹ In-group favoritism refers to the phenomenon that decision makers tend to behave more benevolently toward members of their own group than toward ‘outsiders’. See Goette et al. (2006), Charness et al. (2007) or Fehr et al. (2008) for experimental evidence that in-group effects affect behavior in distributive tasks and in cooperation and coordination games.

indicating that distributional preferences may only be a minor factor in explaining expert sellers' behavior in our field experiment.

Third Degree Price Discrimination. Instead of invoking convex distributional preferences, one could argue that taxi drivers apply price discrimination between low- and high-income passengers. However, the charging decision of the driver is in a sense a dictator's decision. At the end of a ride, a customer can hardly reject payment by arguing that the demanded fare is above his willingness to pay. This applies to high-income and low-income passengers alike, meaning that the income role should not affect the fare for reasons of price discrimination. In fact, this is what we observe for non-local natives and foreign passengers.

High-Income Passengers Suffer Less From Fraud. One might argue that high-income passengers were charged higher fares because taxi drivers expected them to be able to collect a refund of the travel expenses from a company. This could make the high-income passengers more willing to accept fraud (even if they detected it) than low-income passengers, because the latter would have to pay the consequences of the fraud out of their own pockets. Related to this argument is the experimental finding by Gneezy (2005) that subjects cheat more on others by lying more often when the costs imposed on the other person through lying are smaller. If high-income passengers were perceived as suffering less from overtreatment and overcharging – because they either can afford it more easily than low-income passengers or are more likely to get the expenses refunded – then the results of Gneezy (2005) would imply that they are more often cheated upon, which matches the aggregate findings for overtreatment and overcharging.

VI. CONCLUSION

In this paper we have presented the results of a controlled field experiment on the extent and the determinants of fraudulent behavior in the provision of a frequently consumed credence good: taxi rides. The first contribution of this paper to the literature on credence goods markets has been the exact *and* separate measurement of two frequent and serious problems in the provision of credence goods, i.e., of overtreatment and overcharging. Using portable GPS loggers we have been able to keep track of the taxi drivers' routes in the city of Athens, Greece, meter by meter, allowing us to measure the extent of costly overtreatment. By letting a triple of passengers – each in a different role – ask for the same service at practically the same time we were able to control for a variety of unforeseeable factors, such

as traffic jams. Given the data on the exact length of a route, we have then been able to identify the amount of overcharging.

Overall, we have found that 45% of passengers were taken on detours that accounted for at least 5% of the shortest possible route. The overall average was 8%, or roughly 1 km of the average total length of 12.9 km. Overcharging through manipulating fares was observed relatively infrequently in 11% of observations. Recall that overcharging – once detected – is typically much easier to verify than overtreatment, because there are always possible excuses for taking a detour. From this it follows that the potential material costs of being detected (i.e., the risk of being fined or losing one's license) are higher for the taxi driver in the overcharging dimension than the overtreatment dimension. Therefore, it is no surprise that overtreatment is much more frequent than overcharging.

The second main contribution of this paper has been a controlled manipulation of (i) a passenger's familiarity with the city, and thus with the optimal route, (ii) a passenger's presumed information about the taxi tariff system; and (iii) a passenger's income. By doing so, we have been able to study how these factors affect the extent of fraud in the two dimensions overtreatment and overcharging. Consistent with our hypotheses, we have found that passengers who are less likely perceived as knowledgeable about the optimal route are taken on significantly longer detours, while passengers who are less likely regarded as familiar with the (Greek) taxi tariff system are more extensively overcharged. Overall, our findings show that taxi drivers exploit their informational advantage over non-local natives and foreigners with respect to the optimal route by taking them on longer detours. Additionally, they exploit their informational advantage over foreign passengers about the tariff system by overcharging them more frequently and by a higher amount. We have also found an influence of a passenger's income role on fraud, with high-income passengers being taken on longer routes and being overcharged slightly more than low-income passengers. This hints at the possibility that distributional preferences might affect the provision of credence goods, as has been documented in the laboratory experiments of Kerschbamer et al. (2009) and Dulleck et al. (2011). However, the effects of a passenger's income are only partly significant, and hence the influence of distributional preferences seems weak overall.

Summarizing on a more general level, we would like to stress that the lack of knowledge on the consumer's side on the type or quality of the good that generates the largest surplus and on the type of good received is detrimental for honesty in the provision of and charging for credence goods. This is obviously a pervasive phenomenon across several credence goods markets. For example, anecdotal evidence indicates that car owners with

seemingly less knowledge about a necessary repair are more prone to fraudulent behavior by car mechanics than consumers perceived as having more information.²² Another example is provided by Domenighetti et al. (1993), who have shown that with medical services the probability of receiving a largely unnecessary treatment is negatively correlated with the perceived knowledge of the patient about medical issues. Given this evidence and the results from our field experiment, it seems safe to conclude that better informed consumers are less prone to fraudulent behavior from the sellers' side. This finding has practical implications, of course. Conveying to an expert seller the impression of possessing relevant information (be it true or not), or at least refraining from revealing one's lack of information, can alleviate the problems associated with the provision of credence goods. With car repairs, memorizing some technical terms might help, and in the case of medical treatment the existence of a (fictional) doctor in one's family can be the key to an appropriate treatment. In the case of taxi rides, instructing the driver which route to take might be helpful to demonstrate an ability to verify the optimal route. This conveys one's knowledge about the city or a willingness to take action against overtreatment. A related kind of safeguard against fraud in the market for taxi rides is provided by many conference organizers who are concerned about their participants and thus inform them in advance how much a trip from the airport to the venue typically costs. Judged on the basis of our findings this seems a sensible approach to protect participants from the possible downsides of their informational disadvantage in comparison to the expert sellers of the credence good studied in this paper, i.e., taxi rides.

²² See, for instance, the article "Your Wheels. More Women Are Beginning to Take Peek Under the Hood" in the Los Angeles Times, February 23rd 2000.

REFERENCES

- Andreoni, James, and John Miller.** 2002. "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism." *Econometrica* 70(2): 737-753.
- Bekken, Jon-Terje.** 2003. *Taxi Regulation in Europe – Executive Summary*. Institute of Transport Economics, Norway.
- Bolton, Gary E., and Axel Ockenfels.** 2000. "ERC – A Theory of Equity, Reciprocity and Competition." *American Economic Review* 90(1): 166-193.
- Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler.** 1997. "Labor Supply of New York City Cabdrivers: One Day at a Time." *Quarterly Journal of Economics* 112(2): 407-441.
- Charness, Gary, and Martin Dufwenberg.** 2006. "Promises and Partnerships." *Econometrica* 74(6): 1579-1602.
- Charness, Gary, and Matthew Rabin.** 2002. "Understanding Social Preferences with Simple Tests", *Quarterly Journal of Economics* 117(3): 817-869.
- Charness, Gary, Luca Rigotti, and Aldo Rustichini.** 2007. "Individual Behavior and Group Membership." *American Economic Review* 97(4): 1340–1352.
- Cooper, David J., and John H. Kagel.** 2011. "Other Regarding Preferences: A Selective Survey of Experimental Results." In: John H. Kagel and Alvin Roth (eds.), *Handbook of Experimental Economics, Vol. 2*, Princeton, Princeton University Press, forthcoming.
- Cox, James C., Daniel Friedman, and Vjollca Sadiraj.** 2008. "Revealed Altruism." *Econometrica* 76(1): 31–69.
- Crawford, Vincent P., and Juanjuan Meng.** 2011. "New York City Cabdrivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income." *American Economic Review*: Forthcoming.
- Darby, Michael R., and Edi Karni.** 1973. "Free Competition and the Optimal Amount of Fraud." *Journal of Law and Economics* 16(1): 67-88.
- Domenighetti, Gianfranco, Antoine Casabianca, Felix Gutzwiller, and Sebastiano Martinoli.** 1993. "Revisiting the Most Informed Consumers of Surgical Services: The Physician-Patient." *International Journal of Technology Assessment in Health Care* 9(4): 505-513.
- Dulleck, Uwe, and Rudolf Kerschbamer.** 2006. "On Doctors, Mechanics, and Computer Specialists: The Economics of Credence Goods." *Journal of Economic Literature* 44(1): 5-42.

- Dulleck, Uwe, Rudolf Kerschbamer and Matthias Sutter.** 2011. "The Economics of Credence Goods: On the Role of Liability, Verifiability, Reputation and Competition." *American Economic Review* 101(2): 526-555.
- Farber, Henry.** 2005. "Is Tomorrow Another Day? The Labor Supply of New York City Cabdrivers." *Journal of Political Economy* 113(1): 46-82.
- Farber, Henry.** 2008. "Reference-Dependent Preferences and Labor Supply: The Case of New York City Taxi Drivers." *American Economic Review* 98(3): 1069-1082.
- Fehr, Ernst, Helen Bernhard, and Bettina Rockenbach.** 2008. "Egalitarianism in Young Children." *Nature*, 454(28 August): 1079-1084.
- Fehr, Ernst, and Klaus Schmidt.** 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics* 114(3): 817-868.
- Gneezy, Uri.** 2005. "Deception: The Role of Consequences." *American Economic Review* 95(1): 384-394.
- Goette, Lorenz, David Huffman, and Stephan Meier.** 2006. "The Impact of Group Membership on Cooperation and Norm Enforcement: Evidence Using Random Assignment to Real Social Groups." *American Economic Review*, 96(2): 212-216.
- Iizuka, Toshiaki.** 2007. "Experts' Agency Problems: Evidence From the Prescription Drug Market in Japan." *RAND Journal of Economics* 38(3): 844-862.
- Kerschbamer, Rudolf, Matthias Sutter, and Uwe Dulleck.** 2009. "The Impact of Distributional Preferences on (Experimental) Markets for Expert Services." IZA Discussion Paper No. 4647. IZA Bonn.
- Kingsbury, Douglas.** 1968. *Manipulating the amount of information obtained from a person giving directions*. Unpublished Honors Thesis, Department of Social Relations, Harvard University.
- List, John A.** 2006. "Field Experiments. A Bridge Between Lab and Naturally Occurring Data." *Advances in Economic Analysis and Policy* 6: Article 8.
- List, John A. and David R. Reiley.** 2008. "Field Experiments." in: Steven N. Durlauf and Lawrence E. Blume (eds.) *The New Palgrave Dictionary of Economics*. Second edition. Palgrave Macmillan.
- OECD.** 2007. "Taxi Services: Competition and Regulation 2007". Policy Roundtables. Department of Competition Law and Policy. OECD.

- Schaller, Bruce.** 2005. "A Regression Model of the Number of Taxicabs in U.S. Cities." *Journal of Public Transportation* 8(5): 63-78.
- Schneider, Henry.** 2009. "Agency Problems and Reputation in Expert Services: Evidence from Auto Repair." Johnson School Research Paper Series 15-07.
- The New York City Taxicab Fact Book.** 2006. Schaller Consulting [online]. Available from <http://www.schallerconsult.com/taxi/taxifb.pdf> [Accessed February 12, 2011].
- Wolinsky, Asher.** 1993. "Competition in a Market for Informed Experts' Services." *RAND Journal of Economics* 24(3): 380-398.

TABLES

Table 1: Treatments

		passenger's income role		
		<i>low income</i>	<i>high income</i>	total
passenger's information role	<i>local</i>	29	29	58
	<i>non-local native</i>	29	29	58
	<i>foreigner</i>	29	29	58
	total	87	87	174

Table 2: Overtreatment Index
 (normalizing each route by the shortest one in the triple)

passenger's role	<i>low income</i>	<i>high income</i>	<i>total</i>
<i>local</i>	1.01	1.05	1.03
<i>non-local native</i>	1.06	1.10	1.08
<i>foreigner</i>	1.06	1.11	1.08
<i>total</i>	1.05	1.09	1.07

Table 3: Overcharging

[A] Relative frequency of overcharging			
passenger's role	<i>low income</i>	<i>high income</i>	total
<i>local</i>	0.03	0.07	0.05
<i>non-local native</i>	0.00	0.17	0.09
<i>foreigner</i>	0.21	0.17	0.19
total	0.08	0.14	0.11

[B] Amount of overcharging in €, conditional on overcharging			
passenger's role	<i>low income</i>	<i>high income</i>	total
<i>local</i>	0.94	2.43	1.93
<i>non-local native</i>	n.a.	2.39	2.39
<i>foreigner</i>	5.45	6.27	5.82

Table 4: Fare Index

passenger's role	<i>low income</i>	<i>high income</i>	total
<i>local</i>	1.01	1.07	1.04
<i>non-local native</i>	1.07	1.11	1.09
<i>foreigner</i>	1.19	1.18	1.19
total	1.09	1.12	1.10

Table 5: Tobit Regressions

	(1)	(2)	(3)	(4)	(5)
	<i>Overtreatment</i>	<i>Overtreatment</i>	<i>Amount of</i>	<i>Fare</i>	<i>Fare</i>
	<i>Index</i>	<i>Index</i>	<i>overcharging</i>	<i>Index</i>	<i>Index</i>
			(in €)		
$\neg local^{\#}$	0.096*** (0.034)	0.106*** (0.040)	1.978 (2.721)	0.100** (0.049)	0.128** (0.055)
<i>foreign</i> [#]	0.005 (0.024)	0.006 (0.035)	5.148** (2.403)	0.146** (0.056)	0.181** (0.079)
<i>high income</i> [#]	0.061** (0.027)	0.075** (0.0322)	3.588* (2.045)	0.0521 (0.046)	0.115* (0.064)
$\neg local \times$ <i>high income</i>		-0.018 (0.042)			-0.054 (0.064)
<i>foreign</i> \times <i>high income</i>		-0.002 (0.047)			-0.067 (0.095)
<i>driver</i> <i>female</i>	0.012 (0.055)	0.013 (0.056)		-0.102 (0.100)	-0.108 (0.113)
<i>driver age</i>	-0.001 (0.001)	-0.001 (0.001)	0.069 (0.106)	-0.002 (0.002)	-0.002 (0.002)
<i>time of day</i> [#]	-0.004** (0.002)	-0.004** (0.002)	0.548** (0.216)	-0.001 (0.003)	-0.002 (0.003)
<i>constant</i> [#]	1.038*** (0.059)	1.032*** (0.060)	-27.17*** (7.786)	0.978*** (0.110)	0.960*** (0.105)
Prob > F	0.009	0.015	0.024	0.006	0.017

$N = 174$. *, **, *** denotes significance at the 10%, 5%, 1% level respectively.

Overtreatment Index and Fare Index: Tobit regressions, left-censored at 1. Standard errors clustered by set of three simultaneous observations. Overcharged Amount (in €): Tobit regression, left-censored at 0, robust standard errors.

[#] $\neg local$ is a dummy for passengers revealing that they were not familiar with the city (that is, for *non-local native* passengers and *foreign* passengers), *foreign* is a dummy for passengers speaking in English; *high income* is a dummy for passengers in the role of high-income passengers; *time of day* is an ordinal variable that takes on values from 8 to 23 for each hour from 8 a.m. to 11 p.m.

FIGURES

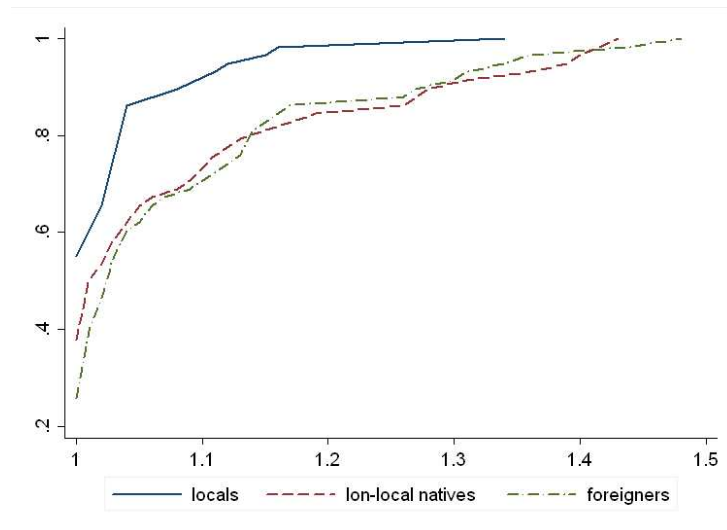


Figure 1: Cumulative Distribution of Overtreatment Index by Passenger's Information Role

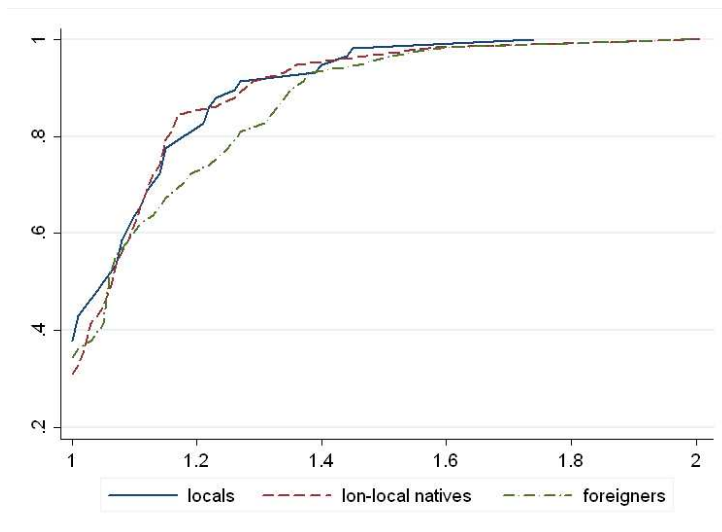


Figure 2: Cumulative Distribution of Duration Index by Passenger's Information Role

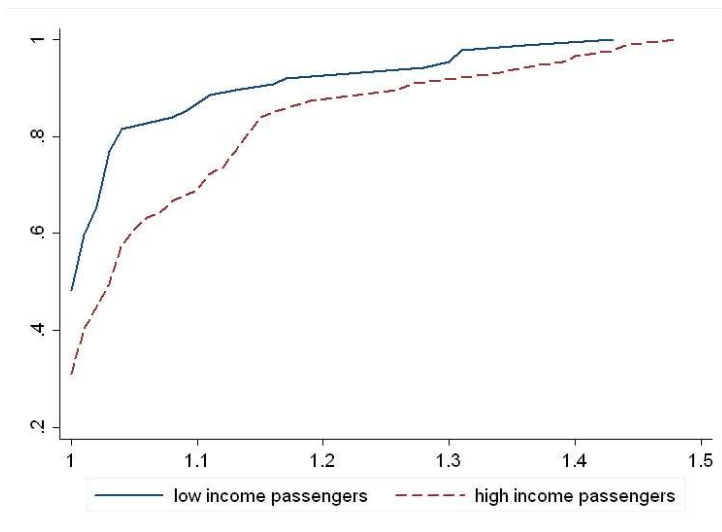


Figure 3: Cumulative Distribution of Overtreatment Index by Passenger's Income Role

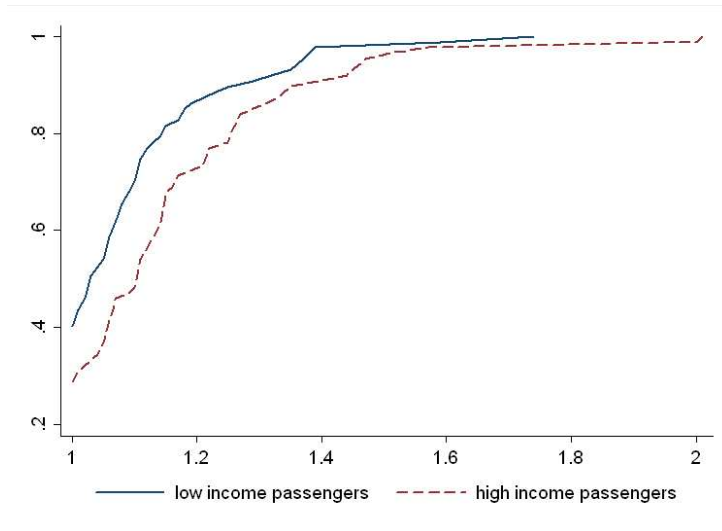


Figure 4: Cumulative Distribution of Duration Index by Passenger's Income Role

APPENDIX

Table A.1: Destinations

Name	Description
Airport	E. Venizelos International Airport
Glyfada	high-income suburb, southern Athens
Karaiskaki Square	run-down neighborhood (central)
Kifissia	high-income residential suburb, northern Athens
Port (Piraeus)	main commercial and touristic port
Syntagma	central square, foreigner area
Train Station	main train station, all intercity trains
Evangelismos	central Athens
Abelokipi	middle-income neighborhood, close to city center
Bus station	main bus station, services mainly to southern and central Greece
Pagрати	central residential area, starting point only

Table A.2: Routes

#	Origin	Destination
1	Pagрати	Port
2	Port	Karaiskaki Square
3	Karaiskaki Square	Kifissia
4	Kifissia	Syntagma
5	Train Station	Port
6	Pagрати	Glyfada
7	Glyfada	Evangelismos
8	Evangelismos	Glyfada
9	Pagрати	Train Station
10	Karaiskaki Square	Abelokipi
11	Abelokipi	Bus Station
12	Bus station	Port
13	Port	Evangelismos
14	Evangelismos	Port
15	Airport	Evangelismos

Table A.3: Alternative Overtreatment Index
(normalizing each route by the shortest possible distance)

passenger's role	<i>low income</i>	<i>high income</i>	total
<i>local</i>	1.03	1.06	1.05
<i>non-local native</i>	1.09	1.12	1.10
<i>foreigner</i>	1.09	1.12	1.10
total	1.07	1.10	1.08

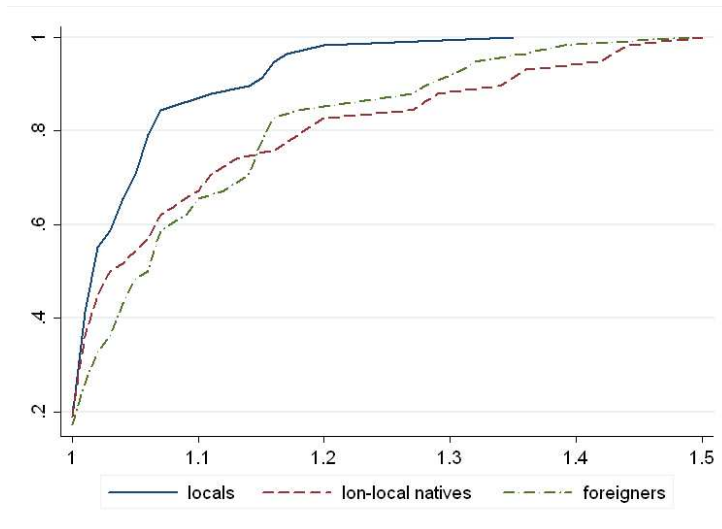


Figure A.1: Cumulative Distribution of Alternative Overtreatment Index by Passenger's Information Role

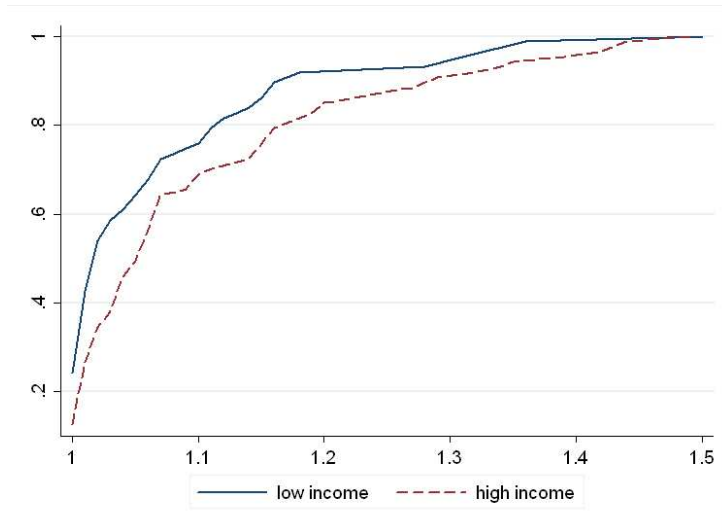
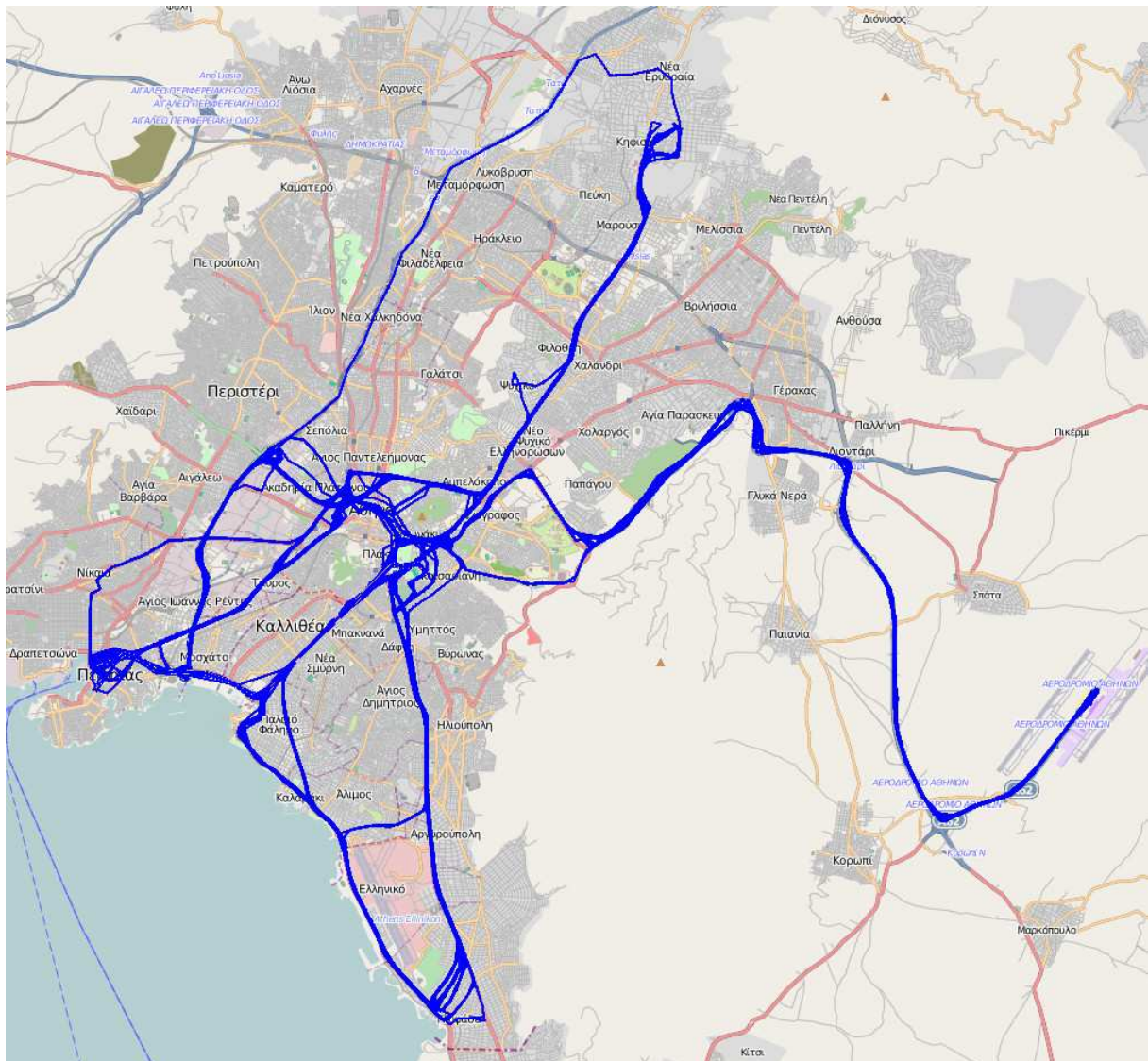


Figure A.2: Cumulative Distribution of Alternative Overtreatment Index by Passenger's Income Role



Picture A.1: GPS satellite logger used during taxi rides



Picture A.2: Routes. © OpenStreetMap contributors, CC-BY-SA. See www.openstreetmap.org and www.creativecommons.org

```

ΕΝΑΡΞΗ ΝΟΜΙΜΗΣ ΑΠΟΔΕΙΞΗΣ
ΑΡ. ΑΠ. ΒΑΡΔΙΑΣ 9
Η/Α 24/07/10 ΧΡ 13:54:26
-----
ΕΚΜΕΤΑΛΛΕΥΤΗΣ:1
ΔΑΓΚΑΛΑΚΗΣ ΔΗΜΗΤΡΙΟΣ
Α.Φ.Μ.:021106526
Δ.Ο.Υ.:20ΓΡΑΦΟΥ
ΑΡ.ΟΧ.:ΤΑΑ 4520
Ν.ΑΤΤΙΚΗΣ
ΤΗΛ.:2108824895
-----
ΣΤΑΘΕΡΑ Κ: 2450
ΣΥΝ. ΔΙΑΝ. ΧΛΜ. 12,00
ΔΙΑΡΚ.ΔΙΑΔΡΟΜΗΣ 00:16:50
-----
ΑΝΑΛΥΣΗ ΔΙΑΔΡΟΜΗΣ
ΤΑΡΙΦΑ-ΑΝΑΜΟΝΗ-ΑΠΟΣΤΑΣΗ
(ΩΡΑ) (ΧΛΜ)
01 00:16:49 12,00
-----
ΑΝΑΛΥΣΗ ΤΙΜΩΝ
ΤΑΡΙΦΑ 01 (11,00%)
ΑΠΟΣΤΑΣΗ 7,86€
ΑΝΑΜΟΝΗ 0,69€
-----
ΣΗΜΑΙΑ(11,00%) 1,16€
ΠΡΟΣΘ.(11,00%) 0,00€
ΕΜΠΕΡΙΕΧ. ΦΠΑ: 0,96€
-----
ΠΛΗΡΩΤΕΟ 9,71€
Α.Μ. ΑΔΗΜΕ:ΤΕΒ 09901162
ΛΗΞΗ ΝΟΜΙΜΗΣ ΑΠΟΔΕΙΞΗΣ

```

Picture A.3: Sample receipt