Gender Differences in Bargaining Outcomes: A Field Experiment on Discrimination

Marco Castillo, Ragan Petrie, Maximo Torero, and Lise Vesterlund

December 2011

Discussion Paper

Interdisciplinary Center for Economic Science
4400 University Drive, MSN 1B2, Fairfax, VA 22030
Tel: +1-703-993-4719 Fax: +1-703-993-4851
ICES Website: http://ices.gmu.edu
ICES RePEc Archive Online at: http://edirc.repec.org/data/icgmuus.html

Gender Differences in Bargaining Outcomes:
A Field Experiment on Discrimination

Marco Castillo*
Ragan Petrie*
Maximo Torero[&]
Lise Vesterlund[#]

December 2011

Abstract: We examine gender differences in bargaining outcomes in a highly competitive and commonly used market: the taxi market in Lima, Peru. Examining the entire path of negotiation we find that men face higher initial prices and rejection rates. These differentials are consistent with both statistical and taste-based discrimination. To identify the source of the inferior treatment of men we conduct an experiment where passengers send a signal on valuation before negotiating. The signal eliminates gender differences and the response is shown only to be consistent with statistical discrimination. Our study secures identification within the market of interest and demonstrates that there are environments where sophisticated statistical inference is the sole source of differential gender outcomes.

* Interdisciplinary Center for Economic Science, George Mason University, [&]International Food Policy Research Institute, [#] University of Pittsburgh and NBER

1. Introduction

Evidence from labor markets (Altonji and Blank 1989, Gunderson 1989, Psacharopoulos and Tzannatos, 1993) and audit studies in a variety of environments (Riach and Rich 2002, Ayres, 1991, Ayres and Siegelman 1995, Heckman 1998, Bertrand and Mullanaithan 2004) show inferior treatment of women and minorities. It is difficult however to determine the extent to which these differences are due to taste (Becker 1975) or to statistical inference on payoff relevant information (Phelps 1972, Arrow 1973) . The main identification problem is that we do not have sufficient data on or knowledge of the information economic agents use to either statistically or taste-based discriminate. The residual found in labor market studies when controlling for productivity relevant factors may result from taste-based discrimination, but it may also result from an inability to control for all relevant productivity differences. A similar problem arises in audit studies where some but not all productivity-relevant characteristics can be observed and manipulated. Heckman (1998) argues that, depending on the nature of these unobservables, audit studies can both identify taste-based discrimination when it does not exist and fail to recognize it when it does. In short, in the absence of direct manipulation of all payoff-relevant information, it is difficult to conclude what model gives rise to remaining differences in market outcomes.[1]

An alternative approach is to manipulate beliefs or directly measure preferences of those engaging in discrimination. For instance, List (2004) finds that agents respond to payoff-relevant information on valuations and perceptions of valuations in a manner consistent with predictions from statistical discrimination. Moreover, Fershtman and Gneezy (2001) and List (2004) show that discriminating agents are not predisposed to engage in taste-based discrimination. Both sets of authors find that transfers in dictator games are not consistent with a distaste for the discriminated group, thus suggesting that differentials are unlikely to be caused by taste-based discrimination.

One concern when assessing taste in say a dictator game is that the differential treatment of interest may be sensitive to the stress and time pressure of the market in which it arises. As suggested by Bertrand, Chugh, and Mullainathan (2005) discrimination may be implicit and sensitive to the cognitive load of the market. Therefore in our design, to determine the source of discrimination, we rely on identification *within* the market of interest. More precisely, after having documented gender differences, we manipulate interactions in the market first to show that information completely can eliminate differences, and second to characterize the types of statistical and taste-based discrimination that are consistent with our data. An attraction of the precise characterization of the possible grounds for discrimination is that we can derive opposing comparative static predictions and directly test

[1] This method has been taken to laboratory experiments (Castillo and Petrie, 2010). Relevant is also Goldin and Rouse (2000) who show that when information on the appearance of music performers is withheld, symphonic orchestras discriminate less against women. Absent information on appearance agents can neither engage in taste-based or statistical discrimination.

competing hypotheses. The paper's contribution lies in demonstrating the advantage of directly identifying the source of discrimination within the market, and in showing that there are markets where sophisticated statistical inference is the sole source of differential outcomes.

Using a group of trained buyers we conduct field experiments to determine whether men and women who use the same, and commonly used, bargaining script receive similar treatment. Holding the behavior of buyers fixed, the emphasis of the study is on examining whether the seller's path of negotiation differs for these two groups of customers, and if so what may explain the observed difference.

For the purpose of the study we examine the market for taxi rides in Lima, Peru. The study was conducted on busy and heavily trafficked routes on weekdays between 8 am – 1 pm. There are a number of reasons why we selected this particular market. First taxi rides account for a sizeable share of household expenditures and the market plays a central role for households in Lima. Second, the channels through which differential outcomes may arise are relatively limited. The market is characterized by free entry and is very competitive. Taxis are widely used and serve as a primary mode of transportation for both genders, causing men and women to have the same experience negotiating for taxi rides. Thus there is limited room for taste-based discrimination and for statistical discrimination on the grounds of experience. Third the market is well suited for experimentation. The full fare of a ride is determined through very short face-to-face negotiations, which can be manipulated in our study. Naturally-occurring negotiations in the market reveal a simple bargaining strategy which is commonly used by both men and women and which is ideally suited for eliciting differential treatment. Specifically many passengers only respond to the driver's price by stating a fixed maximum-acceptable offer. The attraction of this strategy is not only that it is natural to the market and appears to be gender neutral, but also that it allows for a simple script which is easy for our passengers to use and which enables us to elicit the driver's entire path of negotiations.

Mimicking the manner in which negotiations naturally occur in the market, our trained and paid passengers hail a taxi and ask the driver how much it will cost to travel to a predetermined destination. Passengers were only to respond to the driver's offer by stating a predetermined maximum-acceptable offer. This secures that price changes and the ultimate agreement or disagreement only can be done by the driver. Passengers were instructed to complete the transaction in the event that the driver agreed to the maximum-acceptable offer. This assures that the transaction is ended by the driver and not by the passenger's hesitation as agreement is approaching.

The taxi market in Lima is characterized by sequential bargaining and repeated matching. A driver who is matched with a potential passenger engages in sequential bargaining, and in the event of an impasse, both the driver and the passenger return to the market in search of another match. Relying on the work by Samuelson (1992) as well as Fudenberg, Levine and Tirole (1987), we expect that heterogeneous agents will engage in haggling. Drivers screen high-valuation passengers by first quoting a high initial fare and then lowering their price as they become increasingly pessimistic in their assessment of the passenger's valuation, eventually reaching a point of agreement or disagreement. Using a simple theoretical example we demonstrate that the driver, when faced with a passenger who is perceived to be a high-valuation buyer, may use both initial prices and rejection rates to pressure the passenger to accept a high price. Specifically passengers drawn from a distribution with higher valuations may be faced with higher initial prices and rejection rates.

Our results reveal that despite identical bargaining scripts drivers do not treat male and female passengers the same. In sharp contrast to previous studies, the bargaining outcomes for women are superior to those for men. Women are quoted lower initial prices, and conditional on maximum-acceptable offer, they are less likely to be rejected by the driver. The observed gender differences are consistent with drivers perceiving men to be high-valuation passengers and engaging in statistical discrimination. However, these differences are also consistent with drivers having a relative preference for female passengers and engaging in taste-based discrimination. Admittedly the type of taste-based discrimination needed to explain the results is the exact opposite of that needed to explain previously observed gender differences. Nonetheless, as with any taste-based model, it is not difficult to provide examples of preferences that are consistent with the differential treatment.

To identify the source of the gender gap, we conduct a second experiment where passengers, prior to negotiating, send a signal on their valuation. In this second study we take advantage of the fact that it is common for a second taxi to queue up behind an initially hailed taxi to wait for the first negotiation to fail. In the instances where a taxi is waiting behind the first taxi, it is therefore possible to send a signal to the second taxi. Passengers in this second signaling study were instructed to first hail a taxi and ask for the cost to a destination. They were then to visibly reject the first taxi and proceed to negotiate with a second queued taxi, using the maximal-acceptable-offer bargaining script of our initial study. The rejection of the first taxi was intended as a signal of the passenger's low valuation. As the rejection is more costly for a high-valuation passenger this manipulation helps identify the source of the differential treatment. Specifically, if the initial gender differences resulted from statistical discrimination of what is perceived to be a high-valuation male, then the rejection of the first taxi is predicted to reduce the gender difference at the second taxi. By contrast a standard model of taste-based discrimination would predict that the gender gap remains constant. The data from the signaling experiment reveal that the initial prices at the first taxi replicate those of our first study, with men systematically being quoted higher initial prices. However this result changes at the second taxi. We find no gender differences in bargaining outcomes at the second taxi; men and women are quoted the same initial and final price, and they are rejected at the same rate. Whatever the source of the gender differences at the first taxi these differences are eliminated at the second taxi.

While statistical discrimination is consistent with the presence of a gender gap at the first taxi and the absence of one at the second taxi, it is not consistent with taste-based discrimination by both first and second taxis. If taste-based discrimination is the explanation of the initial gender gap then we should see a gender gap at both the first and the second taxi. Nonetheless to fully examine the possibility of taste-based discrimination we also explore the possibility that drivers select to be first rather than second taxis and that only first taxis discriminate based on taste. A closer look at the data reveals that the change in response to gender between the first and second taxi is driven by the first taxi's inferior treatment of men. Men at the first taxi are given a higher initial price and higher rejection rate than at the second taxi, by contrast the treatment of women does not differ between the first and second taxi. Thus the initial gender gap is consistent either with statistical discrimination of high-valuation male passengers or with the first (and not the second) taxi engaging in taste-based discrimination against males.

This more precise characterization of the source of differential treatment allows us to derive opposing comparative static predictions for the two remaining explanations. Testing these we find that behavior only is consistent with statistical discrimination. We conclude that the gender gap in our initial study results from taxi drivers engaging in statistical discrimination of male and female passengers. The second study suggest that the market becomes 'gender blind' when the statistical inferences on men and women are the same.

We begin the paper by characterizing the market of interest. We then present the experimental design and the behavior expected in our study. Our results from the first experiment are presented and followed by a discussion of the possible explanations for the observed differences. We then proceed to the signaling experiment which allows us to determine whether the observed gender differences result from statistical or taste-based discrimination.

2. Why the taxi market in Lima?

There are several reasons why we used the taxi market in Lima, Peru, to conduct our experiment. First, the nature of the negotiation makes it well suited for experimentation. Second, the characteristics of the market reduce the number of channels through which discrimination may work and make the market well suited for examining statistical discrimination. Third, the market plays a central role to households in Lima and the cost of an unsuccessful negotiation is substantial for both passengers and drivers. We justify each of these claims below.

First the nature of negotiation in this market provides us with substantial control over both the negotiated item and the manner in which the negotiation is conducted. The negotiated item is well defined and solely involves agreement on a fare to secure transportation from one point to another.[2] Taxis do not have meters, prices are not fixed by route or zone, and the fare is determined entirely by a face-to-face negotiation. As there is no tipping, the negotiated price captures the entire cost of the fare.

[2] By minimizing the ambiguity about the negotiated item we make it less likely that men and women are expected to achieve different negotiation outcomes (Bowles, Babcock and McGinn, 2005).

Furthermore negotiations are very quick and involve limited verbal exchange. The negotiation is conducted entirely through the window of the taxi while the passenger is still standing on the street. The passenger does not get in the taxi until an agreement has been reached. These characteristics make it easy to manipulate and control the negotiation as we can provide a short, natural, and gender-neutral negotiation script.[3]

Second, the number of factors that may give rise to differential gender outcomes is reduced. Taste-based discrimination is limited by the very competitive nature of the market.[4] The taxi market in Lima is mostly unregulated and there is no limitation on the number of taxis. Legislation was passed in the early 1990's which allowed any person the right to provide public transportation. Combined with a large reduction in public employment and an influx of reconstituted cars, the number of taxis and drivers increased dramatically. While there are taxis on the streets that have gone through licensing by the government, roughly 50 percent of the taxis are unlicensed (JICA, 2005).[5] According to the Metropolitan Transportation Commission there are about 200,000 taxis in the city of Lima. By comparison, New York City has 53,000 licensed taxis and a population of 8.3 million people. Thus with a population in Lima of 7.7 million, the number of taxis is four times larger.[6] Taxis account for an astonishing 28 percent of all motor vehicles on the road and there is approximately one taxi for every 12 working adults in Lima.

An indicator of the fierce competition in the market is that taxi drivers, in hope of a failed initial negotiation, commonly will pull up behind a taxi that is in the process of negotiating with a potential passenger. Taxi drivers spend about half of their time empty and driving around looking for passengers (JICA, 2005). Another indicator is the low earnings of the taxi drivers. Despite working an average of 13 hours per day, a driver's net daily earnings are between 30 and 50 soles which correspond to the minimum daily wage.[7]

While competition reduces the role for taste-based discrimination, the role of statistical discrimination is reduced by passengers (male and female) being very experienced. Taxis are a common mode of transportation for most segments of the population. People take taxis to work, students take them to

---

[3] We note later that an observational study was used to both confirm that naturally occurring negotiations lead to gender differences, and to identify a bargaining script which is commonly used by both men and women.

[4] As noted by Heckman (1998) taste-based discrimination is only eliminated in the long run if the supply of drivers is perfectly elastic at zero price. While perfect competition does not eliminate taste-based discrimination, the driver's small daily earnings make it less likely that they are willing to sacrifice payment to engage in taste-based discrimination.

[5] Estimates on the level of informality in the market vary. According to *El Comercio* of November 6, 2010, the number of illegal taxis in Lima is 32%.

[6] This number is much larger than what a city this size can handle. In 2007, the President of the Peruvian association of drivers (Fechop) estimated the optimal number of taxis should be about 25,000.

[7] There are two types of taxi drivers: those that own their car and those that rent. Approximately half of the authorized drivers own their car. For those that rent, drivers pay a fixed rental fee, between 30 and 60 soles ($10-$20) per day, and they pay their own gas and then keep the money from the fares. The types of cars used for taxis are quite diverse. It is common to see Korean TICOs and Japanese sedans and station wagons some of which have had their steering wheel switched from right to left. Our passengers were instructed not to hail the small and less expensive TICOs.

school and parents take them to go shopping or to drop off children (JICA, 2005). Taxis are used by a representative sample of the population and are used by both men and women. This is in part a reflection of the fact that many people do not own a car in Lima. According to a national survey (Encuesta Nacional de Hogares, 2009), only 17.8% of households in Metropolitan Lima have an automobile, and many of these households use them for business purposes rather than as a mode of transportation.[8] With taxis providing an average of 20 rides per day the estimated number of rides per day is 3 to 4 million. According to the 2007 Peruvian census, metropolitan Lima has about 1.8 million households. Assuming that a taxi passenger takes two rides per day this corresponds to one member of every household taking a taxi every day.

Since taxis are widely used, and the current conditions of the market have been in place for at least 15 years, it is reasonable to argue that this is a market where both male and female passengers are experienced. Thus the role of experienced-based statistical discrimination is limited relative to previously examined markets (e.g., car and sports card markets).[9] That being said, drivers are also very experienced and it is not unlikely that they engage in statistical discrimination to extract rent from high-valuation passengers.[10]

Finally this is a market that plays a central role to households in Lima. Despite each negotiated fare being relatively small, the stakes involved in the negotiations are large. Transportation is an important component of the consumption basket for households. The Peruvian Institute of Statistics finds that an average household in Lima spends about 8.8 percent of their monthly budget on transportation services. According to the IMF, the average per capita monthly income in Peru is about 950 soles. Hence a person who spends 5 soles in taxis daily will consume about 16 percent of their monthly income on taxis alone. This suggests that passengers have a substantial incentive to bargain for the best possible price. The same holds for the taxi driver, as each successful negotiation represents between 5-7 percent of their daily income.

In sum, the Lima taxi market is an attractive environment for examining whether professional male taxi drivers, despite the competitive market and the substantial experience of the passengers, differ in their treatment of male and female passengers.[11] In examining a market where the sources of differential treatment are limited, we are better able to identify the extent to which drivers engage in statistical discrimination to extract revenue from high-valuation passengers.

---

[8] By comparison 46 percent of households in New York City own cars (www.streetsblog.org/2011/04/06/new-yorks-car-ownership-rate-is-on-the-rise/)

[9] See Ayres (1991), Ayres and Siegelman (1995) and List (2004).

[10] Being a taxi driver is the main occupation for 90 percent of drivers (JICA, 2005).

[11] Two recent studies also examine behavior in the taxi market. Keniston (2011) examines the market for local autorickshaw transportation in Jaipur, India, to identify the welfare implications of bargaining relative to giving passengers the option of a fixed price. Balafoutas et al (2011) instead examine the taxi market in Greece to study fraud.

3. Theory

The salient feature of the market under study is sequential bargaining and repeated matching. Buyers and sellers meet, negotiate sequentially until they reach an agreement, or return to the market in search of a suitable match.[12] Samuelson (1992) shows that in a market where bargainers are matched randomly it is possible that disagreement occurs even when gains from trade are common knowledge. The reason is that the negotiating parties may reach a point where the return from a new match exceeds that of the existing match.[13] Fudenberg, Levine and Tirole (1987) demonstrate that haggling, when it is possible to bargain with a sequence of agents, occurs only when transaction costs are sufficiently high. Intuitively, when the seller's cost of finding a new buyer is low, the seller prefers negotiating with a new buyer over continued negotiation with a buyer that has revealed a low willingness to pay. While sellers with low switching cost will charge either a low price or a high take-it-or-leave-it price, sellers with high switching cost may engage in haggling. We draw on both of these models to build an example that shows how bargaining and screening may play out in our environment. Using a model of sequential bargaining and random matching, we start with an example of only one population. We then extend the example to two identifiable populations to demonstrate the comparative statics that may be expected from statistical discrimination in our environment.

Consider a market with an infinite number of buyers (denoted B) and sellers (denoted S). At each time, $t = 1,2,\ldots$, buyers and sellers are matched randomly. Sellers are matched with a buyer with probability θ_S ($\theta_S < 1$) and buyers are matched to a seller with probability θ_B ($\theta_B < 1$). Parameters θ_S and θ_B represent the level of friction in this market and capture the transaction cost of reaching a bargaining impasse. Sellers are assumed to produce the good at a cost of zero. Buyers' valuation of the good is either $\overline{v}$ with probability π or $\underline{v}$ with probability $1-\pi$. Buyers' valuation of the good is private information. For the sake of the example, we assume that $1 < \underline{v} < 2 < \overline{v} < 3$ and that prices can only take integer values.[14] The distribution of buyer's values and the cost of production of the seller are common knowledge. Buyers and sellers discount each period according to discount factors δ_S and δ_B with $0 < \delta_i < 1$ for $i = B, S$. If a seller is matched with a buyer, the seller quotes a price p that the buyer can accept or reject. If the buyer accepts, the buyer earns $v - p$ and the seller earns p. Both agents leave the market and are replaced by identical replicas. If the buyer rejects, the seller and then the buyer have to sequentially decide whether to continue negotiating or return to the pool of unmatched agents. Matches are broken if a matched agent decides to leave the match. If agents remain matched, the seller lists a second price that the buyer can either accept or reject. This process is repeated until they either reach an agreement or the match is broken. If they remain matched, but never reach an agreement, both agents receive a payoff of zero.

---

[12] In naturally occurring negotiations in the market we find that three in ten result in disagreement.
[13] The lowest valuation of a buyer is higher than the highest cost of a seller.
[14] In the market under study, all observed prices are integer valued.

In this model agents cannot learn about the agents they have not interacted with and the environment remains stable. Samuelson (1992) characterizes two types of equilibria which may arise depending on the value of π.[15] There is an equilibrium in which sellers list a price of 1. In this equilibrium all buyers accept the price of 1, but reject a price of 2. There is another equilibrium in which sellers make a take-it-or-leave-it offer of 2 that only buyers with value $\overline{v}$ accept. Sellers enforce screening by abandoning negotiations after observing a rejection. Relying on the insights of Fudenberg, Levine, and Tirole (1987), it is also possible to construct an equilibrium where haggling occurs. For instance, there is a haggling equilibrium in which sellers state a price of 2 as soon they are matched with a new buyer. Sellers then randomize between switching to a price of 1 and abandoning the negotiation. This equilibrium requires that sellers are indifferent between listing a price of 1 and leaving the negotiation. This is possible if the price of 2 is accepted by buyers with value $\overline{v}$ with probability $p = \frac{1-\delta_S}{(2-\delta_S)\cdot\pi}$ (see Appendix for derivation). For a buyer with value $\overline{v}$ to be indifferent between accepting a price of 2 and waiting for prices to drop to 1, the seller and buyer must remain matched and the seller must drop prices to 1 with probability $q = \frac{\overline{v}-2-\delta_B\cdot V_B(\overline{v})}{(\overline{v}-1-V_B(\overline{v}))\cdot\delta_B}$, where $V_B(\overline{v})$ is the equilibrium payoff of a buyer with value $\overline{v}$.[16]

The equilibrium with haggling requires that the portion of high-valuation buyers is sufficiently large $\pi > \frac{1-\delta_S}{2-\delta_S}$ and the take-it-or-leave-it equilibrium requires an even larger portion of high-valuation buyers in the market $\pi > \frac{1-\delta_S}{(2-\delta_S)\cdot\theta_S}$. The result is intuitive. Sellers may engage in haggling as a way to screen buyers when the likelihood of finding a high-value buyer is sufficiently high, but they may find it profitable to completely screen out low-valuation buyers if the fraction of high-valuation buyers is even higher.

Next we consider the case where there are two different and identifiable groups of buyers, Group 1 and Group 2. Since the haggling equilibrium requires the seller to be indifferent between offering a price of 1 and stopping the negotiation, both the price and the rate of rejection can be used to statistically discriminate. The seller may reject buyers of one group more often in an attempt to extract larger prices from them. For instance, suppose that members of Group 1 have valuations that take values $\overline{v}_1$ and $\underline{v}$ and that members of Group 2 have valuations than take values $\overline{v}_2$ and $\underline{v}$ only. Suppose further that $1 < \underline{v} < 2 < \overline{v}_2 < 3 < \overline{v}_1 < 4$ and the proportion of high types in each group is high enough for a haggling equilibrium to exist for each group independently. It is clear that, regardless of the proportion of buyers belonging to Group 1 and Group 2, if sellers are randomly matched to members of these

---

[15] We assume that high-valuation buyers prefer to wait for a lower price (i.e., $\overline{v} - 2 < \delta_B \cdot (\overline{v} - 1)$). This allows us to construct the haggling equilibrium below.

[16] Of course it is also possible to obtain a haggling equilibrium if some buyers strictly prefer to obtain the good sooner at a price of 2 rather than waiting for a price of 1. In this case neither buyers nor sellers need to randomize (see e.g., Fudenberg, Levine and Tirole, 1987).

groups there will be an equilibrium of the game where sellers start negotiation with Group 1 at price 3 and start negotiations with Group 2 at price 2. Moreover, since a mixing equilibrium requires that $\overline{v}_1 - 3 = (\overline{v}_1 - 1) \cdot \delta_S \cdot q_1 + V_B(\overline{v}_1) \cdot (1 - q_1)$ and $\overline{v}_2 - 2 = (\overline{v}_2 - 1) \cdot \delta_S \cdot q_2 + V_B(\overline{v}_2) \cdot (1 - q_2)$, where q_1 and q_2 are the probabilities that prices drop to 1 for Group 1 and Group 2, respectively, it follows that it is possible that in equilibrium $q_1 < q_2$. Thus the high-valuation Group 1 may face higher prices and a lower probability that the negotiation continues and prices drop in the next period (i.e., they face higher rejection rates).

Finally, note that the model also allows for differential treatment when the distribution of values across populations is identical but the members of Group 1 and Group 2 differ in their discount factors or cost of switching to a new match. Members for whom delay is more costly are more likely to be rejected and less likely to receive a discount. Differential treatment will be possible even if sellers are trying to extract the same set of prices from both groups. Thus conditional on receiving the same prices members of the different groups may face different rejection rates.[17]

In sum, haggling is likely to occur as sellers use prices and rejection rates to extract higher prices from high-value buyers. High-value buyers may face higher initial offers and rejection rates. Moreover, differences in rejection rates may persist even when buyers face the same sequence of price offers.

4. Experimental Design

The protocol of our experiment is as follows. Six men and six women are trained to be 'taxi passengers.' They negotiate for a predetermined taxi fare to travel from one destination to another. Passengers are instructed to negotiate for and travel along a number of different routes. Each route consists of three locations, that is, they travel from destination A to B, then from B to C and then return from C to A. Passengers travel in the reverse direction as well.

At each location they hail a taxi at random, approach the passenger window and ask: "How much would it cost to go to X?".[18] After the taxi driver quotes a price, the 'passenger' counters with an experimenter determined price p_{max}. We will refer to p_{max} as the "maximum-acceptable offer." The specific response to the taxi driver's price is nothing other than "p_{max}." There is no other dialogue. The passenger repeats this response after each price from the taxi driver and continues until either the taxi driver accepts the price or leaves in disagreement. Thus the negotiations can only be terminated by the driver. If the first price quoted by the taxi driver is p_{max}, then there is no need to negotiate. If the driver accepts the

[17] The equilibrium with haggling can also be extended to allow for heterogeneity of sellers' cost of production.

[18] To secure comparable taxis we instructed passengers to only hail taxis and not to approach taxis that were waiting. They were also asked only to hail sedans and station wagons., whereas Ticos (very small, light-weight cars) were to be avoided as they generally get a lower price.

passenger's offer, the passenger takes the taxi to the predetermined destination. If the offer is rejected and the taxi leaves in disagreement, the passenger is instructed to step away from the street, take out a cell phone as if they just received a call, and wait for the street to clear of any taxis that might have seen that the passenger's last negotiation ended unsuccessfully. From the passenger's perspective, any subsequent negations can therefore be perceived as fresh. Passengers who, after a period of time, failed to reach the predetermined price for travel to a location were instructed to take a taxi (at a possibly higher price) to the next location.[19] This protocol permits us to obtain observations for all passengers on all routes across the morning.

Our negotiation script limits the responses one may expect to see from passengers in the market. The leftmost panel in Figure 1 presents the structure of a standard negotiation. Taxi drivers state a price after learning the intended destination. Passengers then decide whether to accept this price, continue negotiations and make a counteroffer or leave the negotiations altogether. After this, taxi drivers decide whether to accept the counteroffer, leave or make a counter-counter-offer. Our design is based on this standard procedure with the modification that only taxi drivers end the negotiation and our passengers are instructed to solely state a maximum-acceptable offer "p_{max}". The bargaining structure of our study is represented in the rightmost panel of Figure 1.

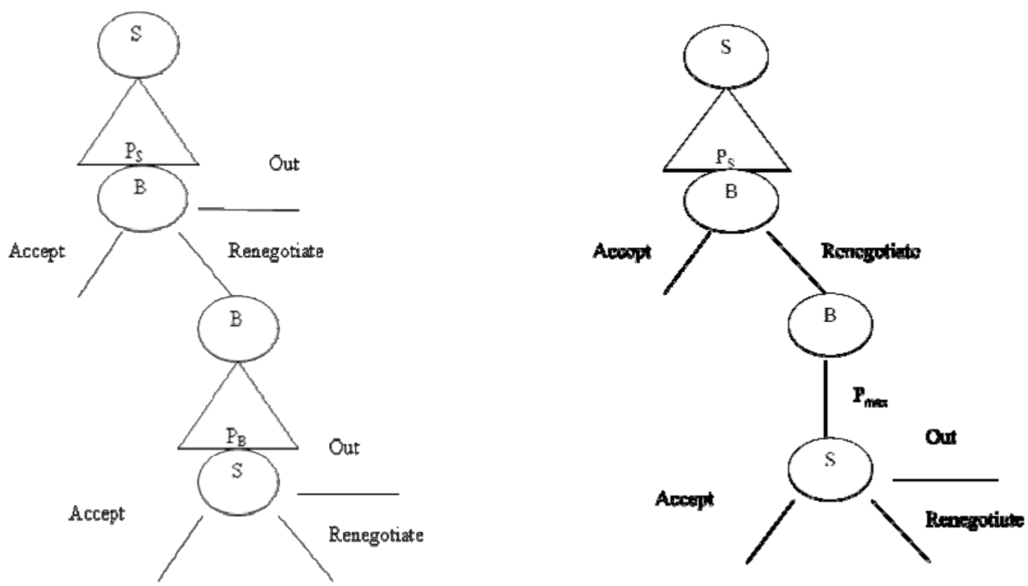

Figure 1: Structure of Bargaining in the Taxi Market

[19] These observations are not included in our data analysis.

There are a number of reasons why we selected this very simple bargaining structure. First, the bargaining approach is similar to that used in the market. It is common practice for customers to approach the passenger-side window and ask for a price with the expectation that some negotiation will ensue. Second, while there is substantial variation in how passengers respond to the driver's stated price in the market, the strategy of simply responding with a maximum-acceptable offer is commonly used.

To design the bargaining script in the experiment, we conducted a study where we observed which strategies passengers naturally use in the market. We recorded 113 of these negotiations. Of these we found that at least 15 percent of passengers only respond with a maximum-acceptable offer. Importantly men and women were equally likely to use the strategy, with 17 percent of women and 14 percent of men employing such a strategy. Note however that the reported frequencies likely are lower bound estimates of how common the strategy is. The reason is that we can only identify the strategy when we observe passengers responding to the driver more than once. Specifically we can only classify a bargaining strategy as being one of stating a maximum-acceptable offer when the passenger states the same offer repeatedly and does it at least twice. If we condition on negotiations that lasted two or more rounds, we find that passengers were observed using the 'maximum-acceptable-offer strategy' 49 percent of the time, with the number being 50 percent for men and 47 percent for women. Thus the bargaining script used in our study is commonly observed in the market and appears to be gender neutral.[20]

A third attraction of the bargaining script is that it has a very limited language, and thus make it easier for our passengers to follow the instructions of staying as neutral as possible, avoiding facial expression and intonation. In longer and more complicated negotiations it is not only difficult to derive a gender-neutral bargaining script, but also to secure that no other information is revealed in the process of the negotiation.[21]

Finally, the script along with passengers completing the transaction allows us to observe the entire path of negotiation. Importantly we are able to determine whether the driver's initial price is the same for men and women, whether the response to gender differs over the course of the negotiation and ultimately whether men and women are equally likely to be rejected by the driver.

---

[20] As suggested by Goldberg (1996) it may be problematic to hold the bargaining script fixed if different classes of buyers do not act the same. There are a number of reasons why this is less of a concern in our study. In addition to finding that that gender differences in initial and final price arise in naturally occurring negotiations in the market, we also find that our short and common gender-neutral bargaining script only results in a gender gap at the first taxi. Our second signaling study shows that these initial gender differences are eliminated when passengers use the same bargaining script, but first send a signal on valuation.

[21] To better control for non-verbal cues in longer negotiations researchers have begun to instead rely on pre-recorded and carefully scripted negotiations, see e.g., Bowles, Babcock, and Lei (2007).

The other parameters of the study were chosen as follows. The ‘maximum-acceptable offers’ used for the study were selected to be low enough that they would trigger negotiations and possibly rejections. Consultation with several taxi drivers and taxi companies helped us select ‘maximum-acceptable offers’ which were close to the lowest price a taxi driver would accept for a particular route.

To secure a competitive environment, we conduct the experiment at central and busy business locations between 8 am to 1 pm, Monday through Friday. Focusing on these hours helps secure that the objective of travel is comparable for men and women and that the drivers have similar outside options.[22] We examine negotiations on 30 different routes over a total of 9 days. The distance between the three points on each route varied greatly. The two shortest routes were 1.2 miles and 1.6 miles long and the two longest routes were 3.8 miles and 3.9 miles long. We chose several routes and distances to ensure that the results are robust and not simply a reflection of a particular population of taxi drivers favoring certain routes.

The six men and six women passengers were chosen so that we have ‘couples’ for whom the primary difference is gender.[23] The two members of a couple are chosen to have similar age, appearance and height. All passengers are trained in the same way and by the same experimenters. All passengers dress alike to avoid attire that might signal personal characteristics. In particular, all participants wear dark pants and a plain, long-sleeve, dark shirt for the entire period of the study. Women do not wear make-up, and men are clean shaven. Neither of these characteristics differ from common attire or appearance in the market. Our trained passengers are paid 15 soles for transportation to the study and another 45 soles per day to conduct the study.

Each passenger carries a small notebook to keep a record of the prices of the negotiation, the time of the negotiation, car characteristics, market conditions etc. In addition they also carry an MP3 player that is used to record negotiations.[24] The recordings of the negotiations allow us to verify the passengers’ recorded data and to reconstruct data in case of faulty note taking. The recordings also serve as a monitor that the passenger follows the experimental protocol. We verified the recordings, and all passengers followed the protocol.

5. Results

We begin by broadly characterizing the negotiations we observed in the market. Drivers are expected to use prices and rejection rates to screen high-valuation passengers. That is statistical discrimination is

[22] Our analysis nonetheless controls for route, time and direction of travel.

[23] Ultimately our second signaling study will help substantiate this claim. While we find significant and robust gender differences in the first study, the response to a signal on valuations cause drivers to treat male and female passengers the same at the second taxi. The path of negotiation seen at the second taxi appears to be gender blind.

[24] The MP3 player has an external microphone that is clipped to the passenger’s shirt or pocket and looks like the passenger listens to music. As in the U.S., many people in Peru walk around and listen to music on MP3 players.

expected to result in high initial prices which decrease as the driver becomes more pessimistic in the assessment of the passenger's valuation. Ultimately a subset of drivers is expected to reject the passenger because the maximum-acceptable offer falls below the driver's outside option or because the driver reaches a point at which he is indifferent between accepting or rejecting the passenger. Having characterized the path of negotiation we proceed to investigate whether it is sensitive to the passenger's gender. If drivers perceive male and female passengers as having different value distributions, then we may find that they are given different initial prices and face different rejection rates. Statistical discrimination would be consistent with high-valuation passengers facing higher initial prices and rejection rates.

5.1. Basic Results

Our data consists of 1,090 negotiations between a male driver and one of our 12 trained passengers.[25] As expected the competition in the market was fierce. Our passengers reported that for 70 percent of the negotiations a second taxi pulled up behind the first taxi to wait for the first negotiation to fail. Interestingly the driver of the first taxi did not respond to a taxi waiting immediately behind him. The driver's initial price, second price and the rate of rejection do not respond to the presence of the second taxi.[26] This absence of a response is perhaps an indication of both the passenger's transaction cost in moving from one taxi to the next and of the high concentration of taxis. Whether a taxi was waiting behind for a failed negotiation, alternate taxis were readily available in the vicinity of the locations we examined.

We begin by determining whether we succeeded in selecting maximum-acceptable offers which were low enough to trigger negotiations and potentially rejections. Conditional on the route's maximum-acceptable offer, Table 1 reports the distribution of initial prices quoted by drivers. Note first that drivers only stated prices in integers. Prices never included cents. Second, the driver's initial prices were never below the maximum-acceptable offer we used on the route, and only rarely (2.8 percent) did the two prices coincide. Conditional on the maximum-acceptable offer, we see substantial variation in the prices passengers initially received. The heterogeneity in initial prices is to be expected given that we are considering different routes at each price, and that the market conditions and the driver's outside option are changing over the course of the day. The mode of the initial price is highlighted for each maximum-acceptable offer, and shows that the gap between the driver's first price and the passenger's maximum-acceptable offer was most commonly 2 soles. An initial price differential of 2 soles was seen in 38 percent of the negotiations, and this was the modal price differential for routes with maximum-

[25] All drivers in our study were male. To our knowledge, there are no female drivers in this taxi market.

[26] A dummy indicating that there is a second waiting taxi during the negotiation does not significantly affect the initial offer, second offer or rejection rate. For initial offer the coefficient on a second taxi dummy is 0.03 (p-value = 0.67), for the second offer it is -0.12 (p-value = 0.34), and for the probability of rejection it is 0.06 (p-value = 0.15). All regressions include dummies for time, route and date, and standard errors are clustered on the passenger.

acceptable offers of 3, 4, and 6 soles. The modal initial price differential was however 3 soles on routes with a maximum-acceptable offer of 5. Thus these routes appear to have been priced more aggressively than our other routes. Overall 40 percent of our negotiations start off with a 3 soles or greater gap between the driver's first price and the maximum-acceptable offer.

Table 1: Distribution of Initial Prices by Maximum-Acceptable Offer

	Maximum-acceptable offer				
Initial price	3	4	5	6	Total
3	7	0	0	0	7
4	18	9	0	0	27
5	**51**	100	9	0	160
6	20	**212**	31	5	268
7	19	128	51	66	264
8	3	57	**107**	**98**	265
9	1	5	21	20	47
10	0	4	24	20	48
12	0	0	1	1	2
13	0	0	0	1	1
15	0	0	0	1	1
Total	119	515	244	212	1,090
Average Price (s.d.)	5.3 (1.2)	6.3 (1.0)	7.7 (1.2)	8.0 (1.1)	6.8 (1.4)
Rejection Rate	55.5%	63.9%	73.4%	50.5%	62.5%

Note: The highlighted bold entries indicate the modal price for each maximum-acceptable offer.

Importantly the maximum-acceptable offers we chose were large enough to trigger rejections. The last row in Table 1 reports the rate at which the passenger and driver failed to reach an agreement during the negotiation. Sixty two percent of the negotiations ended with the driver's final price being higher than the maximum-acceptable offer and the driver rejecting the passenger. The rate of rejection was largest on the 5-sole routes where the initial-price differential was largest.

In describing the high rate of rejection in our study, it is important to note that rejections are common in this market. For comparison, we observed taxi negotiations at four of the locations and during the hours we examined in our study. The objective was to determine the rejection rates naturally seen in the market. Although we do not have information on prices or the pattern of negotiations, we can observe whether a negotiation ends with the passenger getting in the taxi or moving on to negotiate with a different taxi. Of the 211 negotiations we observed, 156 were negotiations when the passenger first entered the market (i.e., negotiating with the first taxi). We found that 28 percent of these new negotiations and 29 percent of all negotiations failed. While this rate of failure is smaller than that observed in our study, it nonetheless makes clear that rejections are common in the market.

Our data reveal substantial differences in the length of the negotiation. Table 2 presents the outcome of the negotiations at each round. Here we define a negotiation round to consist of the driver asking for a price, the passenger responding with the maximum-acceptable offer and the driver deciding whether to accept or reject the passenger. In the first negotiation round (row 1) 20 percent of the passengers were accepted at the maximum-acceptable offer, 28 percent of passengers were rejected, and 52 percent negotiated for an additional round. As the duration of the negotiation increases the rate of rejection increases and it becomes less and less likely that the driver enters another round of negotiations.

Table 2: Distribution of Negotiation Outcomes
(row percentage in parentheses)

	Acceptances	Rejections	Renegotiations	Total
Round 1	191 (20)	303 (28)	566 (52)	1090
Round 2	136 (24)	271 (48)	159 (28)	566
Round 3	44 (28)	90 (57)	25 (16)	159
Round 4	7 (28)	16 (64)	2 (8)	25

While our trained passengers observed substantial variation in the initial price and in the length of the negotiation, changes in driver prices were quite similar over the course of the negotiation. The driver's second price was most likely to be one sol smaller than the first, discounts of two soles or more were only seen in 15 percent of our negotiations. Figure 2 shows the price path for negotiations that lasted 2 or more rounds and those that lasted 3 or more rounds. The graph shows that average prices dropped about one sol from the first to the second round of negotiation and about 40 cents from the second to the third round.

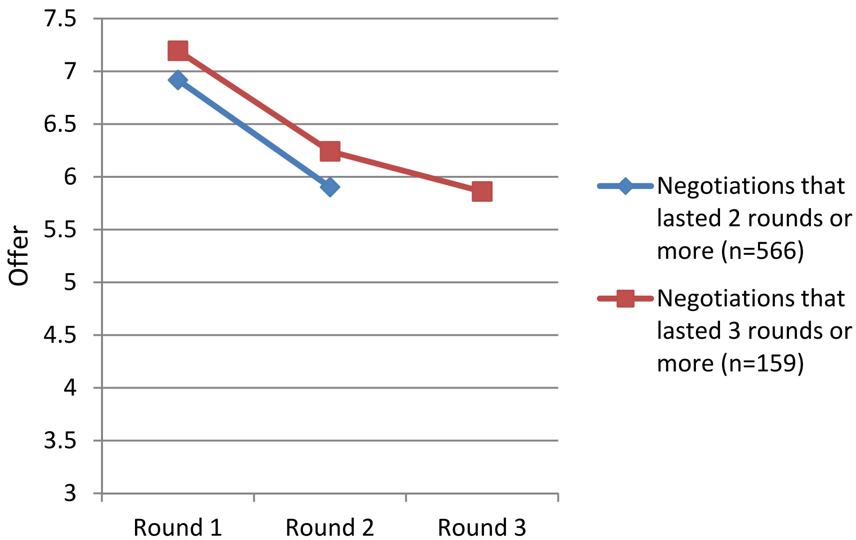

Figure 2: Price Path for Negotiations

As indicated by the greater rejection rate on the more aggressively priced 5-sole routes, the outcome of the negotiation is sensitive to the gap between the driver's initial price and the passenger's maximum-acceptable offer. Figure 3 shows the outcome of the negotiation conditional on the initial price differential. As noted earlier the modal price differential between the initial price and the maximum-acceptable offer was 2 soles. The probability of reaching an agreement decreases substantially with the gap in initial prices. When the initial gap in prices is 2 soles or less, 56.5 percent of negotiations end with the driver accepting the maximum-acceptable offer, however this percentage drops to a mere 8.9 percent when the initial price gap is 3 soles or more.

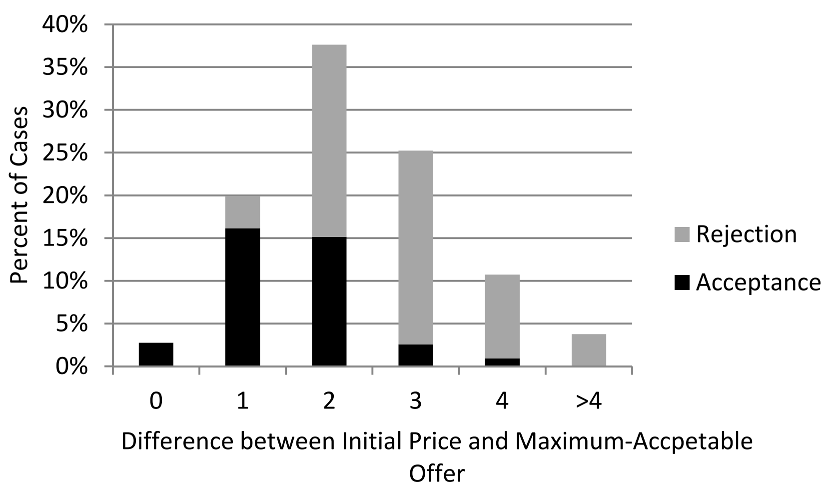

Figure 3: Bargaining Outcome Conditional on Difference Between Initial Price and Maximum-Acceptable Offer

The aggregate data clearly show that the selected maximum-acceptable offers were low enough to trigger both negotiations and rejections. As expected the driver initially asks for a high initial price, then lowers the price and ultimately rejects the passenger when it becomes unlikely that there are gains from trade.

5.2 Gender Differences

Next we examine whether the path of negotiations differ by gender. Recall that our experimental design aimed to minimize differences between men and women. The market is one where men and women are equally experienced; passengers were traveling on identical routes; they were dressed alike; they used the same gender-neutral bargaining script; and the driver's outside option of rejecting a passenger was independent of gender.

We begin by examining whether the driver's initial prices varied by gender. In sharp contrast to previous studies (e.g., Ayres, 1991; Ayres and Siegelman, 1995; and List, 2004), we find that men systematically are quoted higher prices than women. Controlling for the route, direction of travel, time of day, day of the week and clustering on the individual passenger, Table 3 column (1) shows that initial prices given to men were 0.23 soles higher than those for women. With an average initial price for women of 6.76 soles this corresponds to a significant gender gap of 3.4 percent.

Table 3: OLS Regressions on Initial and Last Acceptable Price

	(1)	(2)
VARIABLES	Initial price	Final Acceptable Price
Male	0.23	0.35
	(0.03)	(0.03)
Constant	6.71	5.51
	(0.00)	(0.00)
Observations	1090	1090
R-squared	0.06	0.09

p-values in parentheses. Date, time, and route fixed effects. Standard errors clustered by passenger.

The gender gap in prices continues over the course of the negotiation. Table 3 column (2) reports the results for the driver's final acceptable price, that is the final price the driver stated as being acceptable to him (either the last price stated by the driver before he rejected the passenger or the passenger's maximum-acceptable offer in the event that the driver accepted the offer). As seen in column (2) the driver's final acceptable price for male passengers is on average 0.35 soles higher than for female passengers. With an average final price to women of 5.72 this corresponds to a gender gap of 6.1 percent.

Caution however should be used in interpreting this latter result as it does not control for the length and outcome of the negotiation. That is, some of the included negotiations failed while others succeeded. To evaluate the gender gap in prices over the course of the negotiation it is helpful to compare prices conditional on the negotiation reaching a particular negotiation round. Table 4 reports the gender gap in the first, second and third price the driver stated. The results for the first price are precisely those in Table 3 column (1). We see that the gender gap in prices remains over the course of a negotiation, with men consistently receiving prices which are at least 0.2 soles greater than those received by women. The persistent gender gap suggests that the slopes of the price curves are similar, and that the differential price results from the intercept of the price curve being higher for men than for women.[27]

Table 4: OLS Regressions on Prices across Rounds by Gender[28]

	(1)	(2)	(3)
VARIABLES	Initial Price	Second Price	Third Price
Male	0.23	0.28	0.30
	(0.03)	(0.04)	(0.03)
Constant	6.71	5.96	5.70
	(0.00)	(0.00)	(0.00)
Observations	1090	566	159
R-squared	0.06	0.08	0.13

p-values in parentheses. Date, time, and route fixed effects. Standard errors clustered by passenger.

Next we examine whether there are gender differences in the likelihood by which a negotiation ends in a rejection. We find that the rate at which negotiations fail is 68.5 percent for men and 55.3 percent for women. Conditioning on the maximum-acceptable offer we report the rejection rates by gender in Figure 4.[29] Independent of the maximum-acceptable offer, negotiations with female passengers are more likely to succeed. A 20 percentage point gender gap in rejection rates is seen for all but the 5-soles routes. The most likely explanation for the smaller gender gap on the 5-sole routes is that these routes were very aggressively priced and hence neither men nor women were likely to be accepted on these routes.

[27] Of course selection makes comparisons of the coefficients difficult. If we instead look at the 566 negotiations that lasted two rounds or more we find that the gender gap on the first offer is 0.25 and 0.28 on the second offer.

[28] Because men are rejected more frequently than women, the women in our study are spending more time riding taxis which in turn implies that a larger fraction of our negotiations are done by men. Using routes as the unit of observation, we have that the average proportion of observations produced by men is 55.7 percent (sd 17.4, median 55).

[29] The results are based on matching estimators. For each negotiation with a man we identify a comparable negotiation with a woman. The matches are exact regarding day, route, hour, and maximum-acceptable offer. Note that this is a more strict comparison between male and female rejection rates than a regression since it does not require parametric assumptions.

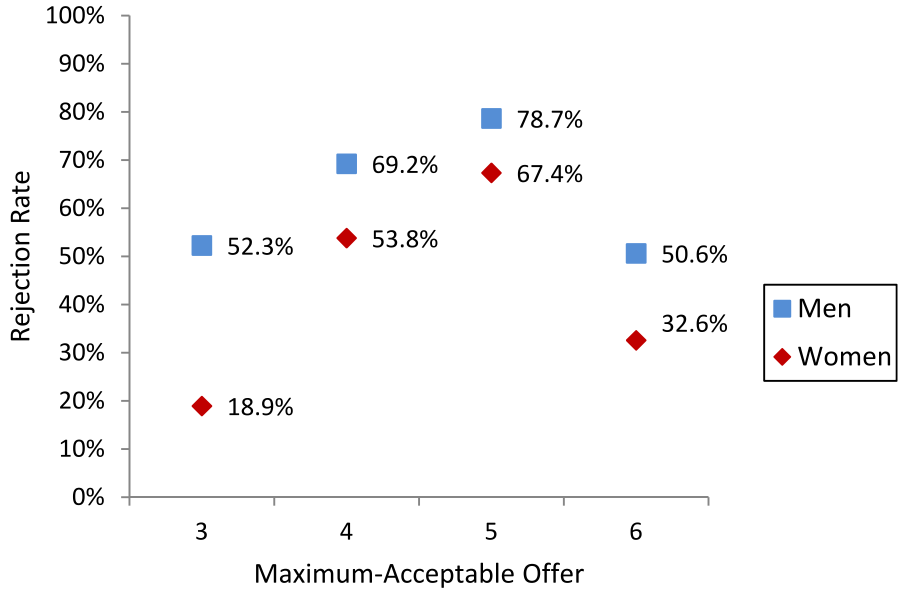

Figure 4. Rejection Rate by Maximum-Acceptable Offer (using matching methods)

Table 5 reports the corresponding OLS regression when we cluster on the individual passenger and include fixed-effects for routes and direction of traffic and controls for day of the week and hour of the day. The overall rejection rate for men is 11 percentage points greater than for women.

Table 5: OLS Regressions on Rejection Rate by Gender

VARIABLES	Rejection Rate
Male	0.11
	(0.06)
Constant	0.49
	(0.00)
Observations	1090
R-squared	0.04

p-values in parentheses. Date, time, and route fixed effects. Standard errors clustered by passenger.

For simplicity we opt to present our results using linear regression analysis, however the results continue to hold when using other methods (e.g., conditional logit and count models). The results on initial price and rejections are also robust to controls for the quality of the taxi and for whether another taxi is waiting behind for the negotiation to fail.[30] As a further test of robustness we find that the results are not driven by a particular couple of trained passengers. Recall that the experiment was designed such that the six men and women in our study could be placed into six pairs of a comparable male and female. Our results are robust to controlling for each of these couples as well as to us randomly dropping any one couple from the analysis.[31]

In summary we find that women are given a significant, substantial, and robust advantage over men. Not only do women face lower initial prices, but they maintain this price advantage as the negotiation progresses. Ultimately the driver is more likely to accept the female than the male passenger. Next we discuss what might give rise to these differences.

6. Explaining the Gender Difference in Bargaining Outcomes

As suggested by our analysis in section 3 we find that drivers do not treat the two types of passengers the same. Consistent with drivers perceiving men to be high-valuation passengers, males are quoted higher initial prices and are ultimately rejected more frequently than females.[32] Finding that behavior is consistent with statistical discrimination however does not enable us to infer that this is what drives the differential treatment in our study. The results may just as well be explained by taste-based discrimination. The higher initial prices and higher rejection rates for males are also consistent with drivers ranking the transportation of female passengers above that of males, be it because they benefit from having female passengers or because they find it costly to have male passengers.

In arguing that taste-based discrimination may explain the advantage given to women, it is interesting to note that the preference differential needed to explain our results is precisely the opposite of that seen in the previous studies where women are placed at a disadvantage relative to men. Nonetheless it is not difficult to provide explanations for why drivers may have a relative preference for female passengers. It may be that women are perceived as being in greater need of a ride and that the driver out of altruism, empathy, or chivalry is more inclined to offer a ride to a female passenger. Or it may be that the driver prefers the company of women and therefore is inclined to accept their low price offers. Alternatively the distaste for male passengers may result from the aggressive negotiation by male

[30] When controlling for a taxi waiting behind, the coefficients on the gender gap for initial price and rejection do not change and the coefficient on the taxi waiting behind is insignificant. Coefficient on taxi behind is -0.00 (p-value=0.99) for initial offer and 0.05 (p-value=0.23) for rejection. Controlling for the quality of the taxi (old, new or normal) also does not change the coefficients on the gender gap for initial price and rejection.

[31] The gender gap in initial price remains at 0.23 (p-value=0.02) if we add dummies for each couple, and the gap in rejection rates is estimated at 0.12 (p-value=0.01). We get similar estimates of the gender gap if we instead drop one couple at a time from the analysis.

[32] Note that we cannot rule out the possibility that the higher rejection of men is caused by drivers viewing a low offer from a perceived high-valuation passenger as being unfair. Central to this argument however is that the differential response is due to the passenger's perceived valuation.

passengers being viewed as more offensive, and thus causing the driver not to give into the male demand for lower prices. Indeed laboratory experiments find evidence in line with this type of behavior. For example, Eckel and Grossman (2001) study ultimatum game bargaining and find that offers coming from women are more likely to be accepted than when the same offer comes from a man. Sutter, Bosman, Kocher, and van Winden (2009) examine a power-to-take game, where one person first takes a share of a pie, and the second person then shrinks the pie.[33] Relative to mixed-gender pairing they find that men paired with men claim a larger share of the pie and retaliate more by shrinking the pie.[34]

While we will directly test the role of taste-based discrimination, one may question whether there is much room for it in this very competitive market. Certainly examination of our data provides no evidence that drivers are willing to sacrifice earnings to either secure a ride with a female passenger or prevent one from a male passenger. For example, if the driver benefits from having a female passenger then that benefit should be increasing with the length of the requested ride. Conditional on maximum-acceptable offer we would therefore expect to see the gender gap in rejections to increase with the length of the ride. Figure 4 suggests that this is not the case as the gender gap in rejection is not larger for rides with higher maximum-acceptable offers which tend to be longer. The same result is seen when instead we condition on the actual length of the ride. Examining routes that were longer than the median route in our study we find, if anything, that the gender gap is smaller on longer routes.[35]

Similar evidence is seen when we look at the time the driver spent transporting passengers. If drivers enjoy the company of women, then they may be able to extend the benefit by selecting a slightly longer route when driving with female passengers. However regression analysis shows no significant difference in the duration of the trip. The average duration of a trip for females was 12.08 min (sd 6.34 min) and 12.14 min (sd. 5.41 min) for males, this difference is not significant (p-value>0.8).[36]

---

[33] I.e., a convex ultimatum game equivalent to Rabin's squishy game. See Andreoni, Castillo and Petrie (2003) for additional evidence on the convex ultimatum game.

[34] The male drivers may see agreeing to the low price as losing a competition, and this may cause them to reject the male passenger at a higher rate. Laboratory as well as field studies document that men are more eager to compete than women (see e.g., Niederle and Vesterlund, 2010, 2011, Croson and Gneezy 2009 for reviews). If a driver views the negotiation with a male as a two-person competition then he may be more reluctant to give into the demands as doing so may be seen as losing the competition.

[35] Clustering on passenger and controlling for date and time fixed effects, a regression of rejection rate generates coefficients of 0.19 (p-value 0.039) on male, 0.06 (0.35) on a dummy indicating whether the trip was above median length, and -0.11 (0.200) on a male and length interaction term. Controlling for maximum-acceptable offer changes the coefficient on the interaction term to -0.12 (0.142). While insignificant the coefficients on the interaction terms indicate that the gender gap in rejection rates decreases rather than increases for long trips.

[36] Another reason to prefer female passengers may be that male passengers are perceived as more dangerous, or that the driver thinks that it is more likely that a male passenger will leave without paying the taxi fare. There are a number of reasons why this is not a likely explanation for our results. First, the study was conducted during regular business hours, on very busy and public routes, and all of our passengers were well dressed and groomed. Thus it is unlikely that the passengers were viewed as being dangerous. Second, if the passenger was viewed as dangerous the driver would have been better off not stopping to initially engage in the negotiation. Third, it is unlikely that a passenger who is planning to skip out on paying the fare would

While our ancillary data analysis suggest that it is unlikely that taste-based discrimination is the main source of the observed gender differences, as in previous studies identification is complicated by the fact that both models are consistent with the observed comparative static. To better evaluate the source of the differential gender treatment, we therefore conducted an additional experimental study which helps separate the two hypotheses.

7. Taste-based and Statistical Discrimination: A Signaling Experiment

To understand the role statistical and taste-based discrimination may play in causing the gender gap in bargaining outcomes, we conducted an experiment where passengers send a signal on their valuation to the driver. This second study takes advantage of the fact that second taxis commonly pull up to wait for the negotiation with the first taxi to fail. In the instances where a second taxi is waiting behind the first taxi, it is therefore possible for the passenger to send a signal to the second taxi. We first describe the design and predicted comparative statics of this second (signaling) study. We then examine whether the results are consistent with taste-based and/or statistical discrimination.

7.1. Experimental Design and Predictions

The signaling experiment follows a protocol similar to that in the first study. The main difference is that a passenger interacts with two drivers, instead of one. The procedure is as follows. The passenger hails a taxi at random. When the taxi pulls up, the passenger asks the driver, through the passenger side window, "How much would it cost to go to X?". The driver states a price, and the passenger shakes his head, steps back from the taxi, and lets the taxi leave. The second taxi, waiting behind the first taxi that the passenger just rejected, pulls up to the passenger. The passenger approaches the taxi and asks "How much would it cost to go to X?". The bargaining protocol then proceeds in the same manner as in our first experiment. The procedures of the first and second experiment are the same. The passenger records the details of the negotiation, as well as the driver and vehicle characteristics of the first and second taxi.

The reason the passenger steps back from the first taxi and shakes his head is to show the second taxi that it is the passenger that rejected the first taxi (not the first taxi rejecting the passenger). This is the common manner in which taxis are rejected by passengers. In the event that no second taxi is waiting behind the first taxi (this happens around 25% of the time), the passenger rejects the first taxi, takes out his cell phone, as if he received a call, and steps away from the street to let the traffic clear. Once any traffic that might have seen the previous negotiation clears, the passenger starts the process again.

---

engage in negotiations in the first place. Finally, and most importantly the absence of a gender gap in the second signaling study demonstrates that the potential ‘concern’ for male passengers is eliminated at the second taxi.

Twelve passengers (6 men and 6 women) participated in a total of 488 negotiations for the signaling experiment. The routes, time of day and maximum-acceptable offers were similar to those used in the first study. To secure that we have comparable observations across the first experiment and the signaling experiment, a subset of the observations were collected by having passengers alternate between the instructions for the signaling study and the first study. This secures that we have observations for the same individuals under both protocols and with the same market conditions.

What changes in initial prices and rejection rates do we anticipate at the second taxi? If the signal at the first taxi causes a significantly large decrease in the second driver's assessment of the passenger's valuation then we should see lower initial prices at the second taxi. What about the rejection rate? As noted earlier, the rejection rate is determined by the drivers' outside option, the passenger's perceived type and the passenger's maximum-acceptable offer. A driver should only wait for a passenger's bargaining impasse with the first taxi if he expects to reach an agreement with the passenger. This implies that the waiting taxi on average should have a lower outside option than the first taxi. Whether this results in a lower rejection rate at the second taxi depends on the extent to which indifferent drivers alter their rate of rejection.[37]

7.2. Results

For the portion of the study where passengers alternate between the study 1 and study 2 protocols, the data from the study 1 protocol confirm that men and women do not receive the same treatment. Male passengers are quoted higher initial prices and they are more likely to be rejected by the driver. Interestingly these gender differences disappear when the passenger rejects the first taxi and proceeds to the second taxi. Looking solely at the study 2 data, we see in Table 6 column (1) that the first taxi quote men initial prices which are 0.32 soles greater than that those given to women. With a mean initial price for women of 6.66 soles this corresponds to a gender gap in initial prices of 4.8 percent. As seen in column (2), this gender gap in initial prices disappears at the second taxi. The coefficient on male is essentially zero and it is far from significant.

[37] The rate of rejection is predicted to decrease at the second taxi if all drivers (first and second) accept passengers they are indifferent towards accepting.

Table 6: OLS Regressions on Initial Price and Rejection Rate (Study 2)

VARIABLES	(1) Initial Price Taxi 1	(2) Initial Price Taxi 2	(3) Rejection Rate Taxi 2
Male	0.32	-0.01	0.04
	(0.01)	(0.94)	(0.37)
Constant	6.74	6.76	0.56
	(0.00)	(0.00)	(0.00)
Observations	488	488	488
R-squared	0.05	0.03	0.02

p-values in parentheses. Date, time, and route fixed effects. Standard errors clustered by passenger.

The gender gap in rejection rates is also eliminated at the second taxi. In the initial study 1 data we found that the rejection rate for men was 68.5 percent while that of women was 55.5 percent. At the second taxi we instead find a rejection rate of 61.5 percent for men and 59.7 percent for women, thus the rate of rejection does not differ by gender at the second taxi (p=0.69).[38] Table 6 column (3) shows that this result is robust to controls for date, time and route fixed effects.

Whatever the source of the gender differences at the first taxi, these differences are eliminated at the second taxi. This change in behavior sheds light on what may explain the initial gender gap. While the response is consistent with statistical discrimination, it makes it questionable that the study-1 gender gap resulted from taste-based discrimination. Certainly the results are not consistent with both first and second taxi drivers engaging in taste-based discrimination. To fully account for the role of taste-based discrimination we do however recognize that it is consistent with a model where there is selection and only the first taxi engages in taste-based discrimination. While we find no evidence that the observable characteristics of first and second taxis differ, elimination of the explanation that only first taxis engage in taste-based discrimination calls for further analysis.[39] Specifically we characterize the nature of discrimination needed to explain our data. Given this characterization we then derive comparative static predictions from the two competing explanations. Fortunately the two explanations give rise to opposing comparative static predictions which we can use to identify the source of the initial difference.

[38] The more comparable rejection rates may be seen when passengers alternate between the two designs. For the study 1 part of the alternating protocols the rejection rate for men is 75.4 and the rejection rate for women is 59.4. For the study 2 part of the alternating protocols the rejection rate for men is 62.4 and the rejection rate for women is 58.0.

[39] The observable characteristics of the first and second taxis are very similar. The average age of the first taxi driver as assessed by our passengers is 38.3 years and that of the second is 39.6. Thirty-three percent of the time the first taxi is a sedan (40% for the second taxi), and 91% of the time the first taxi is either classified as an average aged or new car (88% of the time for the second taxi). Finally, 93% of the time the first taxi is classified as clean in appearance (94% for the second taxi).

We first determine whether the differential response to gender at the first and second taxi is consistent with the first taxi having a taste for women or a distaste for men. Table 7 column (1) reports on the study-2 data and shows that initial prices at the first taxi are 0.30 soles higher for men than for women, and that the initial price to men at the second taxi decreases by precisely this amount, resulting in there being no gender gap in initial prices at the second taxi.[40] Thus moving from the first to the second taxi decreases the initial price to men, but it does not affect the initial price to women. Because we do not observe rejection rates at the first taxi it is not possible to determine how it differs from that of the second taxi. However we can use our alternating protocol design to find comparable rejection rates for study 1 and for study 2. Looking at the alternating protocol data Table 7 column (2) shows that men in the study-1 design face a rejection rate which is 18 percentage points higher than that for women, however this rejection rate decreases when negotiating with the second taxi in study 2. The rejection rate for men is 13 percentage points lower at the second taxi than it was at a comparable first taxi. The net result is that the rejection rates for women are the same at the first and second taxi, and that the gender gap in rejections is absent at the second taxi.[41] Thus for both the initial price and the rejection rate we find that the lack of a gender gap at the second taxi is caused by the treatment of men being more favorable at the second than at the first taxi. This demonstrates that if taste-based discrimination by the first taxi is driving the differential treatment, then it is caused by the first taxi discriminating against men (rather than in favor of women).

Table 7: OLS Regressions on Change in Initial Price from Taxi 1 to Taxi 2 and Rejection Rate between Study 1 & 2

	(1)	(2)
VARIABLES	Initial price Study 2	Rejection rate Alternating btw study 1&2
Taxi 2	0.05	0.04
	(0.46)	(0.24)
Male	0.30	0.18
	(0.00)	(0.01)
Male x taxi 2	-0.30	-0.13
	(0.01)	(0.05)
Constant	6.78	0.54
	(0.00)	(0.00)
Observations	976	547
R-squared	0.03	0.03

p-values in parentheses. Date, time, and route fixed effects. Standard errors clustered by negotiation.

[40] Testing the hypothesis that there is no gender effect (i.e., the sum of the coefficients on men and men x study 2 equals zero) reveals a p-value of 0.97.

[41] Testing the hypothesis that there is no gender effect (i.e., the sum of the coefficients on men and men x study 2 equals zero) reveals a p-value of 0.46.

Of course the differential response to men between the first and second taxi is also consistent with statistical discrimination. If men are perceived to have higher valuations, then a signal to the contrary will cause the second taxi to have a lower assessment of the male passenger's valuation of the ride, potentially causing a decrease in the initial price and in the rate of rejection. To determine whether statistical discrimination or taste-based discrimination against men by the first taxi best explains the data we exploit the fact that the two explanations do not predict the same changes in the gender gap in response to changes in the market.

Not surprisingly the density and characteristics of the passengers change over the course of the morning. The market is most active in the early morning hours when passengers are traveling for work and school.[42] The population of male and female passengers is more homogenous during this period, as they are both rushing to be at a certain place at a certain time and it is likely that high-valuation men and women both are present.[43] Later in the day the market slows down and the population becomes more heterogeneous with some passengers traveling to meet certain appointments while others are traveling under less time pressure. These changes in the market help us derive opposing comparative statics for the two models.

Consider first how the cost of taste-based discrimination changes. The cost of engaging in taste-based discrimination against men is greater later in the morning when there are fewer potential passengers, and the cost is smaller earlier in the morning when demand is higher and there are more high-valuation female passengers in the market.[44] Thus taste-based discrimination against men (by the first taxi) is predicted to generate a larger gender gap in the morning, than later in the day.

By contrast statistical discrimination predicts a smaller gender gap early in the morning and a larger gender gap later in the day as the male and female passengers become more heterogeneous. When male and female passengers are more similar in their value distributions, we would expect a smaller

---

[42] According to a transportation survey (JICA, 2005) 80% of trips made during peak morning hours (7-9am) are trips to work or school (split roughly equally between the two). As an indication of the high demand at 8 am hour we find that rejection rates are at their highest at that time, they are 68% at 8am and 61% the rest of the morning. Furthermore, JICA (2005) reports that the proportion of taxis without a passenger is the lowest in the morning (26 percent 7-10 am, 39 percent from 11am – 2 pm).

[43] To get a sense of the market we had pictures taken of passengers in our observational study. These pictures were ranked by income by 16 taxi drivers. Consistent with the argument above, male and female passengers traveling between 8 and 9 am were perceived to have higher incomes than those traveling at other times of the day. The scores were based on 1-10 Likert scales that have been standardized by rater (16 raters) and then averaged across raters to get a measure for each picture. Throughout the day men and women are equally likely to use taxis. The labor force participation rates of men and women in the markets under study are similar and high. For all the districts included in the study, the percent of the population that is economically active is 61% for men and 44% for women. If we restrict the sample to people aged 20-50 years the proportion who is economically active is 84% for men and 72% for women.

[44] Note that determining the characteristics of the taste-based discrimination is essential for deriving this comparative static. If instead we were finding evidence consistent with first taxis discriminating in favor of females (and securing female passengers) then it would be more expensive to engage in discrimination early in the day and taste-based discrimination would instead predict a smaller gender gap when the market is busy.

gender gap, and this gap should be increasing as the population of passengers becomes more heterogeneous over the course of the morning.

Table 8 reports the gender gap in initial prices for every hour of the morning from study 1.[45] As the sample at any given hour is reduced substantially, the gender gap is generally not significant. However it is interesting to note that while there is a large and systematic gender gap in initial prices at 10 am, 11 am and 12 pm, the difference is much smaller at the 8 am and 9 am time slot. The low gender gap during the morning hours is consistent with our prediction from statistical discrimination and inconsistent with the prediction from taste-based discrimination.[46] Consistent with the 8 am time slot being one of high demand we note that initial prices are higher for this time slot.[47]

Table 8: OLS Regressions on Initial Price over Course of the Morning (Study 1)

VARIABLES	(1) 8am	(2) 9am	(3) 10am	(4) 11am	(5) 12pm
Male	0.09	0.04	0.29	0.26	0.20
	(0.56)	(0.47)	(0.02)	(0.15)	(0.25)
Constant	7.54	6.27	6.49	6.63	6.55
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	221	237	246	268	118
R-squared	0.04	0.05	0.04	0.05	0.03

p-values in parentheses. Date and route fixed effects. Standard errors clustered by passenger.

To further determine whether the gender gap is significantly smaller early in the morning, we focus on routes that are used primarily for commuting, both at peak times (when there are many high-valuation passengers), and at non-peak times when the market is more diverse. Thus, we look at the change in the gender gap at the peak commuting time (8am to 9am) and nonpeak times (10am to noon). According to the JICA (2005) report on traffic volume in various commuting locations across Lima, traffic peaks during the period of 8-9am and drop precipitously during the 9-10am time period. To test the robustness of our results, we use three different criteria to select commuting routes.

The results are presented in Table 9. We regress initial price on gender and a dummy variable for peak time (8am) and interaction terms of peak time and gender, date and route fixed effects. The interaction term on male and peak time tells us if the difference in the gender gap is significantly lower at peak

[45] The mean initial offer is 7.38 (variance 1.77) at 8am, 6.56 (variance 1.99) at 9 am, 6.72 (variance 2.12) at 10 am, 6.71 (variance 1.94) at 11 am, and 6.96 (variance 2.42) at 12 pm.

[46] Pooling the data from study 1 with those from the first taxi in study 2 reveals the same result, with the smallest gender gap occurring between 8 am and 10 am, a large and significant gender gap between 10 am and noon, and finally a slight drop in the gender gap at the noon hour.

[47] By 8am we refer to negotiations that occur between 8:00-8:59am.

times. The main regression is run on three sets of routes. The first set is routes that start and end in the business and residential district Miraflores. Many people commute to and from Miraflores to go to work early in the morning and to come to shop later in the morning as businesses open. The second set of commuting routes is based on a stricter definition and only examines routes that originate in Miraflores. During peak hours these routes represent middle-class workers who need transportation to work, and during nonpeak hours these routes would have a greater diversity as the purpose of travel becomes more diverse when the shops in the area open for business. The final set of routes consists of routes that are determined to be congested and heavily used for commuting by a transportation survey of traffic in Lima in 2005 (JICA, 2005).[48]

Table 9: Gender Gap in Initial Price Comparing 8am to 10-11am (Study 1)
Using interactions of 8am with all fixed effects (date, route)

	(1)	(2)	(3)
VARIABLES	Miraflores	Leaving Miraflores	JICA
8am	0.83	1.21	0.68
	(0.014)	(0.134)	(0.014)
Male	0.36	0.70	0.30
	(0.068)	(0.006)	(0.122)
Male x 8am	-0.33	-1.08	-0.33
	(0.049)	(0.000)	(0.035)
Constant	7.52	3.77	5.23
	(0.000)	(0.000)	(0.000)
Observations	412	219	385
R-squared	0.46	0.46	0.56
Number of Routes	11	6	14

Date, route fixed effects, errors clustered by passenger. Miraflores = all routes that leave from or go to Miraflores. Leaving Miraflores = all routes that leave from Miraflores. JICA = all routes that are classified as being congested by the JICA, 2005 report.

As can be seen in Table 9, independent of how we define commuting routes, the gender gap in initial prices is significantly smaller during peak hours (8am-9am) compared to nonpeak hours (10am-12pm). When we restrict the sample to routes that are defined to be commuting routes under various criteria, we get a robust result of a smaller gender gap during times when the market is more homogeneous. While the results from the two studies may be seen as evidence that the initial gender gap results from a rather unusual model of taste-based discrimination a more careful analysis ultimately rules out this

[48] The routes used in this classification have volume of roughly 2500+ cars passing in the area during the 8-9am period.

possibility. We find that the differential bargaining outcomes observed in the first study are caused by statistical discrimination.[49]

8. Conclusion

Substantial research has been done to understand differential outcomes by gender in markets, particularly the underlying causes of statistical or taste-based discrimination. Identifying discrimination is tricky because, using only observable differences, it is difficult to derive comparative static predictions from a taste-based model which differ from those of a statistical model. Hence identification often requires that beliefs and tastes be measured. Elicitation of such measures is however complicated by the fact that discrimination may be implicit and thus influenced by the cognitive load of the market. Thus, proper control may require that beliefs and tastes are measured in the market transaction of interest. Recognizing the difficulty associated with such an exercise, we instead design two experiments which enable us to characterize the discrimination needed to explain the results and to derive opposing and testable comparative statics from the two models.

To examine the role of statistical discrimination, we examine a market where differences in valuation are likely to play a central role. In studying the taxi market in Lima, Peru, we select an environment where alternate channels for differential outcomes are likely to be limited. In particular, the market is very competitive and all agents are experienced. We find that drivers are far from gender blind when presented with trained passengers who use the same bargaining script. In sharp contrast to previous studies, we find that men receive worse bargaining outcomes than those received by women. While these differences are consistent with statistical discrimination, it is not possible to eliminate the possibility that differences result from taste-based discrimination. Admittedly the type of taste-based discrimination needed to explain the results is the exact opposite of that needed to explain previously observed gender differences. Nonetheless, as with any taste-based model, it is not difficult to provide examples of preferences that are consistent with the differential treatment.

To disentangle the statistical and taste-based explanations, we therefore conduct an additional signaling experiment. Passengers in this study begin by rejecting a first taxi to send a signal of low valuation to a second (waiting) taxi which they then negotiate with. Despite passengers otherwise using an identical bargaining script, we find that negotiated outcomes at the second taxi are gender blind. The second taxi treats men and women the same. While this response is easily reconciled with statistical discrimination, that is not the case for taste-based discrimination. If the initial results were due to a taste-based bias, then this bias is not present at the second taxi. Thus the signaling result eliminates the possibility that a bias is expressed by all drivers in the market. The case under which we can reconcile the observed differences with taste-based discrimination is when the following three conditions hold. First, there is

[49] As further evidence of statistical discrimination, we find that the effect of moving from the first to the second taxi is largest at 8 am and then decreases over the course of the morning.

selection between drivers who become first and second taxis, second, only those who become first taxis engage in taste-based discrimination, and third, discrimination is driven by distaste for male passengers. That is taste-based discrimination is only consistent with our data if first taxis (and not second taxis) discriminate against men.

The attraction of characterizing the type of taste-based discrimination needed to explain the outcomes is that it enables us to derive opposing comparative statics for the two competing explanations. In doing so we are able to conclude that differential outcomes in this market result from statistical rather than taste-based discrimination.

One interpretation of our finding is that drivers in this market do not have preferences for the gender of the passenger that they are transporting. Another interpretation is that the competitive pressure and associated low earnings prevent drivers from expressing any bias they may have against a certain type of passenger. While Heckman (1998) makes clear that perfect competition does not eliminate the possibility that agents engage in taste-based discrimination, the very low earnings of this market may make it unlikely that drivers can afford to cater to such biases. Consistent with the standard interpretation of Becker (1975) we do not find evidence that taste-based discrimination occurs in this very competitive market.

Throughout the paper we have made clear that we selected the market for this study precisely because the characteristics were such that the channels for differential outcomes were reduced. While we anticipate that the results of this study will hold for markets with similar characteristics, it should not be a surprise if that is not the case for markets with different characteristics. Failure to document taste-based discrimination in this market does not imply that it does not occur in others. The competitive pressure and the earnings of the agents involved are likely to play a central role in markets where biases may be expressed, potentially making some markets more susceptible than others.

What we have shown is first that it is possible to use experimental techniques to disentangle the source of differential outcomes within a market and second that there are large and important markets where differential outcomes arise from agents engaging in sophisticated statistical discrimination. While competitive pressure does not eliminate the possibility that taste-based discrimination arises, it does appear that the low incomes that result from such competition make it unlikely that agents will sacrifice earnings to cater to such biases. Our results suggest that the market under study becomes gender blind when the statistical inference on men and women becomes the same.

References

Andreoni, James, Marco Castillo and Ragan Petrie, “What Do Bargainers’ Preferences Look Like? Experiments with a Convex Ultimatum Game,” *American Economic Review*, 93(3), 2003, pp. 672-85.

Altonji, Joseph G., and Rebecca Blank, “Race and Gender in the Labor Market,” in *Handbook of Labor Economics,* Vol. 3C, O. Ashenfelter and D. Card, eds. (Amsterdam, The Netherlands, Elsevier Science B.V., 1999), pp. 3143–3259.

Arrow, Kenneth, “The Theory of Discrimination,” in *Discrimination in Labor Markets,* O. Ashenfelter and A. Rees, eds. (Princeton, NJ: Princeton University Press, 1972).

Ayres, Ian, “Fair Driving: Gender and Race Discrimination in Retail Car Negotiations,” *Harvard Law Review* 817 (1991), 104(4), pp. 817-72.

Ayres, Ian, and Peter Siegelman, “Gender and Race Discrimination in Bargaining for a New Car,” *American Economic Review,* 85, 1995, pp. 304–321.

Balafoutas, Loukas, Adrian Beck, Rudolf Kerschbamer and Matthias Sutter, “What Drives Taxi Drivers? A Field Experiment on Fraud in a Market for Credence Goods,” University of Innsbruck, working paper. April 2011.

Becker, Gary, *The Economics of Discrimination*, University of Chicago Press, 1975, 2nd edition.

Bertrand, Marianne, Dolly Chugh and Sendhil Mullainathan, “Implicit Discrimination” *The American Economic Review*, 94(2), pp. 94-98.

Bertrand, Marianne and Sendhil Mullainathan, “Are Emily and Greg More Employable than Lakisha and Jamal?” *The American Economic Review*, 94(4), pp. 991-1013.

Bowles, Hannah, Linda Babcock and Linda McGinn, “Constraints and triggers: Situational Mechanics of gender in negotiation,” *Journal of Personality and Social Psychology*, 89(6), 2005, pp. 951-965.

Bowles, Hannah, Linda Babcock and Lai Lei, “Social Incentives for Gender Differences in the Propensity to Initiate Negotiations: Sometimes it does hurt to ask,” *Organizational Behavior and Human Decision Processes*, 103(1), 2007, pp. 84-103.

Castillo, Marco and Ragan Petrie, “Discrimination in the Lab: Does Information Trump Appearance?” *Games and Economic Behavior*, 62(1), 2010, pp. 50-59.

Croson, R. and Gneezy, U. (2009) “Gender Differences in Preferences,” *Journal of Economic Literature*, 47 (2), 1-27.

Eckel, C. and P. Grossman, “Chivalry and Solidarity in Ultimatum Games,” *Economic Inquiry*, 39(2), 2001, 171-188.

Fershtman, Chaim, and Uri Gneezy, “Discrimination in a Segmented Society: An Experimental Approach,” *Quarterly Journal of Economics,* 106, 2001, pp. 351–377.

Fudenberg, D., D. Levine and J. Tirole, "Incomplete Information Bargaining with Outside Opportunities," *Quarterly Journal of Economics*, 102(1), 37-50, 1987.

Goldberg, Pinelopi K., “Dealer price discrimination in new car purchases: Evidence from the consumer expenditure survey,” *Journal of Political Economy* 104, 622–654, 1996

Goldin, Claudia, and Cecilia Rouse, “Orchestrating Impartiality: The Impact of ‘Blind’ Auditions on Female Musicians,” *American Economic Review,* 90, 2000, pp. 715–741.

Gunderson, Morley, “Male-Female Wage Differentials and Policy Responses,” *Journal of Economic Literature*, 27(1), 1989, pp. 46-72.

Heckman, James J., “Detecting Discrimination,” *Journal of Economic Perspectives,* 12, 1998, pp. 101–116.

Japan Internation Cooperation Agency (JICA), Transport Council of Lima and Callao, Ministry of Transportation and Communications of the Republic of Peru. *The Master Plan for Lima and Callao Metropolitan Area: Urban Transportation in the Republic of Peru (Phase I). Final Report.* Vol I. August 2005.

Keniston, Daniel, “Bargaining and Welfare: A Dynamic Structural Analysis of the Autorickshaw Market,” Yale University, mimeo. September 2011.

List, J., “The Nature and Extent of Discrimination in the Marketplace: Evidence from the Field,” *Quarterly Journal of Economics*, 119(1), 2004, pp. 49-89.

Niederle, Muriel and Lise Vesterlund, “Explaining the Gender Gap in Math Test Scores: The Role of Competition,” *Journal of Economic Perspectives*, Spring 2010, Vol 24, Number 2, 129-144

Niederle, Muriel and Lise Vesterlund, “Gender and Competition” forthcoming in *Annual Review, 2011*.

Peru. Instituto Nacional de Estadistica e Informatica. Encuesta Nacional de Hogares, 2009. http://www.inei.gob.pe/web/enaho/.

Phelps, Edmund, “The Statistical Theory of Racism and Sexism,” *American Economic Review,* 62, 1972, pp. 659–661.

Psacharopoulos, George and Zafiris Tzannatos, “Economic and Demographic Effects on Working Women in Latin America,” *Journal of Population Economics*, 6(4), 1993, pp. 293-315.

Riach, Peter A., and Judy Rich, “Field Experiments of Discrimination in the Market Place,” *Economic Journal,* 112, 2002, F480–F518.

Rubinstein, A. and A. Wolinsky, “Equilibrium in a Market with Sequential Bargaining,” *Econometrica*, 53(5), 1985, pp. 1133-1150.

Samuelson, L., “Disagreement in Markets with Matching and Bargaining,” *Review of Economic Studies*, 59(1), 1992, pp.177-185.

Sutter, Matthias, Ronal Bosman, Martin Kocher and Frans van Winden, “Gender Pairing and Bargaining-Beware the Same Sex!” *Experimental Economics*, 12(3), 2009, pp. 318-31.

APPENDIX

The appendix provides more detail on the construction of the equilibria included in the theory section. Unmatched buyers and sellers always search for a new match. The actions available to a seller if matched are to offer a price of 1 or 2 and to remain in the match or leave the match to search in the event that the price is rejected. The actions available to a buyer if matched is to accept or reject an offered price and to remain matched or leave the match to search conditional on the seller remaining in the match.

Table A1 reproduces the proposed equilibria in Samuelson (1992). Strategies depend only on whether the agents are matched or not. Samuelson (1992) shows that the *no return* equilibrium exists for all values of π. Denote by V_S the expected equilibrium payoff of a seller. A matched seller will offer a price of 1 if this is better than eventually returning to the pool of unmatched agents and obtaining $\delta_S V_S$. This will be the outcome if players follow the recommended equilibrium strategies since $V_S = \theta_S \cdot 1 + (1-\theta_S) \cdot \delta_S V_S < 1$. Given the strategies, it is also the case that after a rejection, both seller and buyer prefer to remain matched and trade at 1 than going back to the pool of unmatched agents and be matched again with less than certainty. Finally, since we assume that $\overline{v} - 2 < \delta_B \cdot (\overline{v} - 1)$, it follows that high-value buyers are better off rejecting an offer of 2 trading later at a price of 1.

In the *return* equilibrium, sellers offer a price of 2 and abandon negotiations if the buyer rejects the offer. The equilibrium requires that sellers find it profitable to return to the pool of unmatched agents after a rejection and obtain $\delta_S V_S$ rather than offering a price of 1 and obtaining $\delta_S \cdot 1$. It also requires that offering a price of 2 and obtaining an expected payoff of $\pi \cdot 2 + (1-\pi) \cdot \delta_S V_S$ is better than offering a price of 1 that is accepted by everyone. Since the expected payoff of the seller in this equilibrium is $V_S = \theta_S(\pi \cdot 2 + (1-\pi) \cdot \delta_S V_S) + (1-\theta_S) \cdot \delta_S V_S$, it follows that $V_S > 1$ implies that $\pi \cdot 2 + (1-\pi) \cdot \delta_S V_S > 1$. Given these conditions, this equilibrium is possible if $\pi > \frac{1-\delta_S}{2-\delta_S \cdot \theta_S}$.

In addition to these equilibria we also can construct a *haggling* equilibrium as the one discussed in the theory section. Let p be the probability with which a high-value buyer accepts a price of 2 when offered and let q be the probability with which a seller drops the price to 1 if the price of 2 is rejected. A seller is indifferent between dropping the price to 1 and searching for a new partner if V_S equals 1. p is the probability with which a high-value buyer accepts a price of 2 such that V_S equals 1. However, high-value buyers will be indifferent between accepting and rejecting an offer of 2 if $\overline{v} - 2 = \delta_B(\overline{v} - 1) \cdot q + V_b(\overline{v}) \cdot (1-q)$, where $V_b(\overline{v})$ is the equilibrium payoffs of a buyer of value $\overline{v}$. Note that the above equalities imply that $p = \frac{1-\delta_S}{(2-\delta_S) \cdot \pi}$ and $q = \frac{(\overline{v}-2) - V_b(\overline{v}) \cdot \delta_B}{(\overline{v}-1) \cdot \delta_B - V_b(\overline{v}) \cdot \delta_B}$.[50] This implies that a *haggling* equilibrium

[50] In equilibrium, $V_B(\overline{v}) = \frac{\theta_B \cdot (\overline{v}-2)}{1-\delta_S+\theta_B \cdot \delta_S}$.

can occur only if $\overline{v} < \frac{2-\delta_B}{1-\delta_B}$ and $\pi > \frac{1-\delta_S}{2-\delta_S}$. The *haggling* equilibrium can occur when a screening (*return*) equilibrium cannot.

Table A1								
Strategies of matched agents in period t:								
	Seller		Buyer with value $\underline{v}$			Buyer of value $\overline{v}$		
		search if buyer rejects	response to $p = 1$	response to $p = 2$	search if seller stays	response to $p = 1$	response to $p = 2$	search if seller stays
No return	$p = 1$	no	accept	reject	no	accept	reject	no
Return	$p = 2$	yes	accept	reject	no	accept	accept	no
Strategies of unmatched agents in period t: search								