Honesty in the City

Martin Dufwenberg, Maroš Servátka,
Jorge Tarraso & Radovan Vadovič*

February 22, 2021

Abstract

Lab evidence on trust games involves more cooperation than conventional economic theory predicts. We explore whether this pattern extends to a field setting where (much like in a lab) we are able to control for (lack of) repeat-play and reputation: cab drivers in Mexico City. We find a remarkably high degree of trustworthiness, also with price-haggling, which is predicted to reduce trustworthiness.

Keywords: trust, honesty, reciprocity, field experiment, haggling, taxis

JEL Classification: C72, C90, C93, D91

1 Introduction

During the last quarter-century, economists argued that social preferences shape behavior in important ways. Many laboratory experiments were conducted with students who exhibit qualities researchers deem convenient (like

*MD: University of Arizona, University of Gothenburg, CESifo; martind@eller.arizona.edu. MS: MGSM Experimental Economics Laboratory, Macquarie Business School and University of Economics in Bratislava; maros.servatka@mgsm.edu.au. JT: Libretto; jorge.tarraso@gmail.com; RV: Carleton University; radovan.vadovic@carleton.ca. We thank Paul Feldman for excellent research assistance. We also thank David Reiley and several participants in seminars for helpful comments and discussion. Funding was provided by University of Canterbury.

being accessible, smart, and motivated by low stakes) in settings controlled to rule out confounds of repeat-play or reputation. Over time, the lab evidence has amassed and now paints a picture where trust, cooperation, and honesty are abundant but not universal. It is legitimate to wonder about external validity though, and desirable to perform parallel field tests. It may then be a challenge to maintain control regarding repeat-play and reputation. We attempt to overcome that hurdle by conducting an experiment in the highly decentralized market for taxi rides in Mexico City (MC), where the chances of repeated encounters are minuscule.[1]

Our primary concern is trustworthiness. We flag down cabs at point A, ask them to deliver a CD to point B, paying in advance. From the viewpoint of the driver, the situation is analogous to that of the second-mover (the "trustee") in a lab trust game, in particular versions that allow cheap talk before play.[2] In the lab, trustees cooperate with high but not full frequency. We ask if that pattern translates to the streets of MC.[3]

We ran a pilot for that Baseline treatment, planning to condition further research questions and treatments on the nature of the data. If trustworthiness were low, we would have a treatment adding handshakes & promises to the pre-play communication with the driver to test if such enhanced covenants boost trustworthiness. However, trustworthiness was extremely high already in the Baseline, so we scrapped the promises & handshakes treatment. Instead, we ran a Haggling treatment, predicted to instead re-

[1]According to Mexico's "Secretaria de Transportes y Vialidad" in 2011, 102,110 licensed taxis served almost 9 million people in Mexico City, and over 21 million in the larger metropolitan area. The taxis drivers are self-proprietors holding individual licenses (as opposed working for a taxi company) making it virtually impossible to track them down.

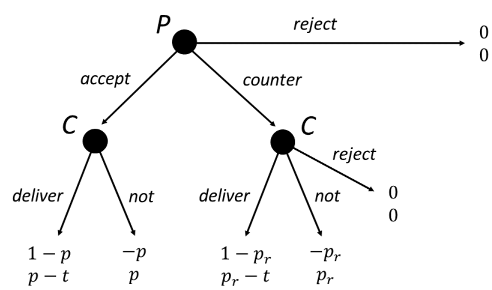
[2]See, e.g., Berg, Dickhaut & McCabe (1995) for a pioneering contribution and Charness & Dufwenberg (2006) and Vanberg (2008) for versions with pre-play communication.

[3]Our experiments were carried out in 2011, before the car-sharing cell-phone technology enabled passengers to track the drivers. A few other recent studies of taxi markets have also addressed questions of cooperation and cheating (Balafoutas, Beck, Kerschbamer & Sutter 2013; Balafoutas, Kerschbamer & Sutter 2017; Castillo, Petrie, Torero & Vesterlund 2013; Bengtsson 2016). In contrast to those studies, our design is geared to generate a direct analog to a trust game in the field.

duce trustworthiness: If a cab driver asked for a price p then we haggled and tried to bring the price down to $p_r < p$. Our hypothesis, which is intuitive but also consistent with reciprocity theory, is that delivery rates will be lower in HAGGLING than in BASELINE.[4]

2 Theory

Assume that cab driver C perceives that he interacts with passenger P as follows:



C has offered to deliver a CD for the price p. In response, P may *reject* (and walk away) or *accept* or *counter* at the lower price $p_r < p$. In the latter two cases, C may choose whether to *deliver* or *not*, and in response to *counter*, C may furthermore also *reject* (and walk away). We normalize payoffs such that each player gets 0 if either player *rejects* and P's value of a safe delivery equals 1. In addition, C faces transportation costs of t. To allow meaningful gains-from-trade, assume that $1 > p > p_r > t > 0$.

If the players are selfish there is a unique subgame perfect equilibrium: C chooses *not* to deliver at each of his nodes and P *rejects* at the root. The outcome is materially inefficient. That conclusion changes if C is sufficiently

[4]Reciprocity has been documented mostly in laboratory experiments (see, e.g., Fehr & Gächter 2000 and Sobel 2005 for discussions).

strongly motivated by kindness-based reciprocity, as modeled by Dufwenberg & Kirchsteiger's (2004) (D&K) notion of sequential reciprocity equilibrium (SRE).[5] To explain, we first describe C's choices as a behavior strategy: Let δ be the probability of *deliver* following *accept*. Let ε be th probability of *deliver* following *counter*. Let ρ be the probability of *reject* following *counter*. In D&K's theory there is key parameter $Y_{CP} \geq 0$ measuring the degree to which C gets utility from reciprocation, with higher values meaning higher sensitivity. The following result provides a behavioral foundation for our experimental hypothesis about the effect of haggling:

Proposition 1 (i) *In any SRE it holds that* $\delta \geq \varepsilon \geq \rho = 0$. (ii) *There exist* $h > \ell > 0$ *such that if* $\ell < Y_{CP} < h$ *then in any SRE it holds that* $\delta > \varepsilon$**.**

We shall test the hypothesis that haggling leads to lower delivery rates, and Proposition 1 shows that this is consistent with reciprocity theory (if $\ell < Y_{CP} < h$)**.** The proof requires an introduction of D&K's formalism; we provide that in the Appendix. Here we just offer the following brief interpretation: In SRE, C always perceives accept by P to be at least as kind or kinder than counter, so C's inclination to be kind in return is then reflected in that $\delta \geq \varepsilon$. Moreover, if C is eager to reciprocate kindness, but not too eager, delivery is strictly more likely following accept than counter.

3 Experimental design

We selected a six-mile route along Via Insurgentes, a straight and major traffic artery in MC. A cab ride takes 20-30 minutes and costs about 30 pesos. We employed two research assistants (RA) who were native speakers. The RA at point A ("RA-A") flagged down a cab, and asked the driver to deliver a CD containing a movie clip to his friend ("RA-B") at point B. RA-A asked for the price, explained that the friend at point B had no money, and proposed to pay up front. The next move depended on the treatment.

[5] D&K extend to extensive games ideas about reciprocity pioneered by Rabin (1993).

The cab drivers were assigned to treatments – Baseline or Haggling – in an alternating fashion. In Baseline, RA-A agreed to the offered fare. In Haggling, RA-A made a counteroffer, subtracting 10 pesos from the driver's proposal. If the driver made a counteroffer, RA-A agreed and paid up. If the driver rejected the lower fare, then RA-A tried to get a discount by making successive offers until the driver agreed. (There were no rejections.) After a deal was struck, RA-A thanked the driver, described what the friend looked like, and paid. As the cab drove off, RA-A discretely recorded the data and sent an SMS to RA-B. Once the cab arrived at point B, RA-B collected the CD and thanked the driver. If the driver demanded additional money, RA-B pretended he did not know that the fare has already been paid, paid the driver, and recorded the sum, the license plate, and the time.

4 Results

Seven cabs refused to be hired. Among the others, we count 31 observations in Baseline and 30 in Haggling.[6] For those, the initial price proposals averaged 38.97 in Baseline[7] and 40.00 in Haggling and were not significantly different ($p = 0.546$; Mann-Whitney ranksum test). The final prices in Haggling averaged 32.70 and were significantly lower than initial proposals in either treatment ($p < 0.001$ for both comparisons). There is no statistical difference in the amount of time it took the cabs to deliver the CD between the two treatments ($p = 0.359$).[8]

Our two main findings are as follows:

[6] RA-A felt uncomfortable haggling with drivers who made very low initial offer (≤ 25 Pesos) and assigned such cabs to Baseline. He similarly assigned an unusually high offer (≥ 60 Pesos) to Haggling. To eliminate possible selection bias at the tails of our initial offer distributions we excluded those data points and conduct the analysis on the remaining 61 observations. Including them does not change our main results.

[7] In Baseline, two cab drivers refused to make a proposal, instead asking RA-A for an offer. RA-A guessed an average fare based on traffic conditions. In one case he offered 30, in the other 50. Both offers were accepted (and the observation assigned to Baseline).

[8] The data is described in more detail by Figures 1 and 2 in the Appendix.

- Comparing our MC setting to all those lab experiments with students (cited in the introduction), our measure of trustworthiness (delivery rates) is *much* higher. With the exception of a single cab driver (in BASELINE), every cab driver who agreed to deliver the CD did so.

- The prediction that delivery rates would be lower following price HAGGLING was not supported; see the previous bullet.[9]

Robustness We conducted a robustness check at a different location, along the street Calle Bolivar between Avenida Hidalgo and Eje 6 Sur (see the Appendix for a map). We only ran the HAGGLING treatment with 20 cabs. *All* of them delivered.

Unexpected forms of cheating While there was virtually universal delivery, in twelve cases (seemingly not distinguishable by treatment), cab drivers cheated in other ways. Their gimmicks included telling RA-B that RA-A had not paid; showing up with a (possibly manipulated) meter read, vastly exceeding the amount originally agreed to and asking for a matching top-up; and claiming that some previously not mentioned extra fee applied.

What did MC locals expect? We didn't in our wildest dreams predict the astonishing level of honesty (as regards delivery rates) exhibited by our MC cab drivers. On seeing the data, we got curious whether also MC locals would find the results surprising. Therefore, we conducted a survey with students at Instituto Tecnológico Autónomo de México (ITAM). We used the technique for eliciting shared opinions developed by Houser & Xiao (2011). Many locals seem likeminded to us. We did not find any evidence of consensus between them that all cabs would deliver. See the Appendix for details.

[9] We are not rejecting the predictions of Proposition 1, which shows that our hypothesis of lower delivery rates in HAGGLING is consistent with reciprocity theory (if $\ell < Y_{CP} < h$). But Proposition 1 is also consistent with the actual data (if $Y_{CP} > h$).

5 Concluding remark

While our results surprised us, in retrospect, and in light of another article of ours perhaps we should have known better? Dufwenberg, Servátka & Vadovič (2017) develop a theory of *informal agreements* in which one of two central assumptions is that once a person enters an informal agreement he or she will not renege.[10] When we designed our experiment, we did not have that theory in mind. We were rather thinking in terms of a comparison to lab experiments with students, and whether we could by treatment marginally affect trustworthiness in the directions described in the introduction. However, it appears that (i) our design generates informal agreements that the CDs be delivered, and (ii) The 2017-theory would do a good job at explaining the data, even though our current experiment was never intended to test it.

Appendix

A.1 Price-haggling and reciprocity

Elements of D&K We focus mainly on C's utility which consists of a material component ($\pi_C(\cdot)$) and a reciprocity component. The latter is the product of how kind C believes that P is to him ($\lambda_{CPC}(\cdot)$) and how kind C is to P in return ($\kappa_{CP}(\cdot)$), weighted by C's reciprocity sensitivity parameter $Y_C \geq 0$. So C's utility has the form $\pi_C(\cdot) + Y_C \times \kappa_{CP}(\cdot) \times \lambda_{CPC}(\cdot)$.

In order to measure kindness, we need to consider the beliefs a player holds about the strategy of the other. A_i is i's set of behavioral strategies, $b_{ij} \in A_i$ is the first-order belief of i about j's strategy, and $c_{iji} \in A_i$ is i's second-order belief of about j's belief about i's strategy.[11] $\pi_j(a_i, b_{ij})$ is the (material) payoff i believes he gives to j (computed as if $a_j = b_{ij}$), and i is kind to j if i believes he gives j a relatively high payoff. Formally, i's

[10]The second assumption, less relevant for our purposes here, is that temptations to reneg affect the form of the informal agreements that people strike. For some related work, see also Miettinen (2013).

[11]All beliefs are point-beliefs, assigning probability 1 to whatever is believed.

kindness is computed by comparing $\pi_j(a_i, b_{ij})$ to the average, or "equitable," payoff that i believes that he could give j, given by

$$\pi_j^e(b_{ij}) = \frac{1}{2}\left(\max_{a_i \in A_i} \pi_j(a_i, b_{ij}) + \min_{a_i \in E} \pi_j(a_i, b_{ij})\right),$$

where $E_i \subseteq A_i$ contains those ("efficient") strategies of i that do not for sure lead to Pareto-inferior material outcomes in any history of play.[12]. In our game $a_i \in A_i \backslash E_i$ iff $i = C$ and a_C puts positive probability on the choice *reject* following *renegotiate.*[13]

i's kindness from choosing a_i when holding belief b_{ij}, is defined as

$$\kappa_{ij}\left(a_i, b_{ij}\right) = \pi_j(a_i, b_{ij}) - \pi_j^e(b_{ij})$$

and i's belief about the kindness of j, λ_{iji}, is derived the same way as κ_{ji}, replacing j's strategy a_j by b_{ij} and by replacing b_{ji} by c:

$$\lambda_{jij}\left(b_{ji}, c_{jij}\right) = \pi_j(b_{ji}, c_{jij}) - \pi_j^e(c_{jij}).$$

In SRE, beliefs coincide with the chosen strategies and at every history of play beliefs are updated to be consistent with reaching that history. Furthermore, at all histories choices must be optimal given the beliefs.

Proof of Proposition 1: Recall that δ, ε, and ρ are the probabilities of, respectively, *deliver* following *accept*, *deliver* following *counter*, and *reject* following *counter*. We establish three lemmas around which the proof is then built:

***Lemma 1*:** *In any SRE,* $\rho = 0$. To see this, note that if a_C puts a positive probability on *reject* following *counter*, then it is not an efficient strategy (as defined earlier in this section). It can never be rationally used except as an

[12]Strategies in $A_i \backslash E_i$ are called "inefficient" because they involve Pareto-decreasing "waste" after some history of play.Refer to D&K (pp. 275-7) for a precise definition and elaboration on why the $E_i \subseteq A_i$ feature is important to the theory.

[13]To see this, note that following choice *renegotiate* choice *deliver* gives both P and C higher material payoff than choice *reject*.

unkind response to a co-player believed to be unkind. However, in our game such a use can be ruled out, because if C believes that P is unkind, then it must be better for P to choose *not* (to deliver) than to *reject;* the former choice brings a higher payoff to C as well as a lower payoff to P than the latter choice does. This implies that $\rho = 0$.

Lemma 2**:** *In any SRE,* $\delta \geq \varepsilon$. Assume to the contrary that $\delta < \varepsilon$. P's kindness depends on his choices and on b_{PC} which specifies his beliefs about δ, ε, and ρ. In an SRE these beliefs are correct and it follows that P must be less kind following *counter* than following *accept*. To see this, refer to the definition of $\kappa_{ij}(\cdot)$ above (letting $i = P; j = C$) and note that since $\rho = 0$ (Lemma 1), $p > p_r$, and $\delta < \varepsilon$ we get

$$\begin{aligned}
\pi_C(\textit{counter}, b_{PC}) &= \varepsilon \cdot (p_r - t) + (1 - \varepsilon - \rho) \cdot p_r + \rho \cdot 0 = p_r - \varepsilon \cdot t \\
&< \\
\pi_C(\textit{accept}, b_{PC}) &= \delta \cdot (p - t) + (1 - \delta) \cdot p = p - \delta \cdot t.
\end{aligned}$$

Moreover, since in SRE C's beliefs about P's beliefs are correct, C similarly perceives that P is more kind following *accept* than following *counter*. Now note that the "marginal material impact" of C's choices, on himself as well as on P, is the same following *accept* as following *counter*. Namely, he can incur-or-not "t" for himself, and he can give or deny "1" to P). If $1 > \varepsilon > \delta \geq 0$ then C must be indifferent between *deliver* and *not* following *counter* (otherwise C wouldn't be willing to mix), but based on what we just said about the marginal material impact, it then follows that he must strictly prefer *deliver* to *not* following *accept*. Hence, $\delta = 1$, a contradiction. And if $1 = \varepsilon > \delta \geq 0$, then C must (weakly) prefer *deliver* to *not* following *counter*, but then again (based on what we said about the marginal material impact) it follows that he must strictly prefer *deliver* to *not* following *accept*. Hence, $\delta = 1$, again a contradiction. We conclude that $\delta \geq \varepsilon$.

Lemma 3**:** *In any SRE, if* $0 < \varepsilon < 1$ *then* $\delta > \varepsilon$. We verify this by contradiction. If the implication were false then either $1 > \varepsilon = \delta > 0$ or $1 > \varepsilon > \delta \geq 0$ would hold. Both of these cases can be considered simultaneously. Applying analogous arguments as in Lemma 2, we can conclude that $\pi_C(accept, b_{PC}) > \pi_C(\ counter, b_{PC})$ and that the marginal material impact of C's decision is the same following *accept* as following *counter*. Therefore, C must be indifferent between *deliver* and *not* in the subgame following *counter*, implying that $\delta = 1$ which is a contradiction.

Part (i) of Proposition 1 is established by combining Lemmas 1 and 2.

It remains to prove part (ii) of Proposition 1. It is helpful to first establish necessary conditions on Y_C for the existence of SRE with, respectively, $\delta = \varepsilon = 0$ and $\delta = \varepsilon = 1$. For the former case, due to Lemma 2, we only need to check that $\delta = 0$ maximizes utility for C following *accept*. Plug relevant numbers into the utility expression $\pi_C(\cdot) + Y_C \times \kappa_{CP}(\cdot) \times \lambda_{CPC}(\cdot)$; one sees that $\kappa_{CP}(\cdot)$ equals $-\frac{1}{2}$ or $\frac{1}{2}$ depending on C's choice,[14] while $\lambda_{CPC}(\cdot)$ equals $p - \frac{1}{2} \cdot (p + 0) = \frac{p}{2}$, so the following inequality holds:

$$\underbrace{p + Y_C \cdot (-\frac{1}{2}) \cdot (\frac{p}{2})}_{\text{utility of } \textit{not}} \geq \underbrace{p - t + Y_C \cdot (\frac{1}{2}) \cdot (\frac{p}{2})}_{\text{utility of } \textit{deliver}}$$

The inequality can be re-written as $Y_C \leq 2t$.

For the latter case ($\delta = \varepsilon = 1$), due to Lemma 2, we only need to check that $\varepsilon = 1$ maximizes utility for C following *counter*. Again $\kappa_{CP}(\cdot)$ equals $-\frac{1}{2}$ or $\frac{1}{2}$ depending on C's choice, but $\lambda_{CPC}(\cdot)$ now equals $p_r - t - \frac{1}{2} \cdot ((p-t)+0) = \frac{2p_r - p - t}{2}$, so the following inequality holds:

$$\underbrace{p_r - t + Y_C \cdot (\frac{1}{2}) \cdot (\frac{2p_r - p - t}{2})}_{\text{utility of } \textit{deliver}} \geq \underbrace{p_r + Y_C \cdot (-\frac{1}{2}) \cdot (\frac{2p_r - p - t}{2})}_{\text{utility of } \textit{not}}$$

If $\frac{2p_r - p - t}{2} \leq 0$, then P's counteroffer is not interpreted as kind, and hence the inequality will not hold for any Y_C. If $\frac{2p_r - p - t}{2} > 0$, then we can re-write the inequality as $Y_C \geq \frac{2t}{2p_r - p - t}$. Note that $\frac{2t}{2p_r - p - t} > 2t$.

Now select ℓ and h, with $\ell < h$, such that neither of the necessary conditions for SRE with $\delta = \varepsilon = 0$ or $\delta = \varepsilon = 1$ hold: If $\frac{2p_r - p - t}{2} \leq 0$, select

[14] For *not* we get $\kappa_{CP} = 0 - \pi_C^e(b_{CP})$ and for *deliver* we get $\kappa_{CP} = 1 - \pi_C^e(b_{CP})$, where $\pi_C^e(b_{CP}) = \pi_C^e(accept) = \frac{1}{2}((1-p) + (-p)) = \frac{1}{2}$.

(any) $\ell > 2t$ and then $h > \ell$. If $\frac{2p_r - p - t}{2} > 0$, select $\ell \in (2t, \frac{2t}{2p_r - p - t})$ and then $h \in (\ell, \frac{2t}{2p_r - p - t})$. Suppose that $Y_C \in (\ell, h)$. Some SRE must exist; this follows from D&K's existence theorem. By design of ℓ and h, neither $\delta = \varepsilon = 0$ or $\delta = \varepsilon = 1$ is compatible with SRE. Lemma 2 imposes the constraint that $\delta \geq \varepsilon$. We can group the remaining possibilities into three cases: $\varepsilon = 0 < \delta$ and $0 < \varepsilon \leq \delta < 1$ and $\varepsilon < \delta = 1$. For the first and last case, obviously $\varepsilon < \delta$. For the middle case, $\varepsilon < \delta$ is implied by Lemma 3. Part (ii) of Proposition 1 follows. ■

A.2 Figures

Figure 1 displays CDFs of the final prices by treatment that were accepted by the drivers.

Figure 1: Distributions of initial offers and final prices

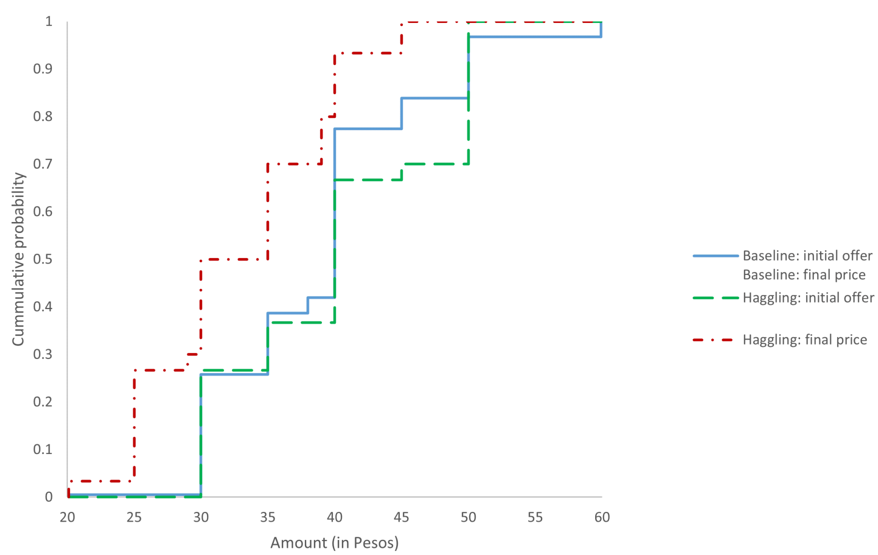
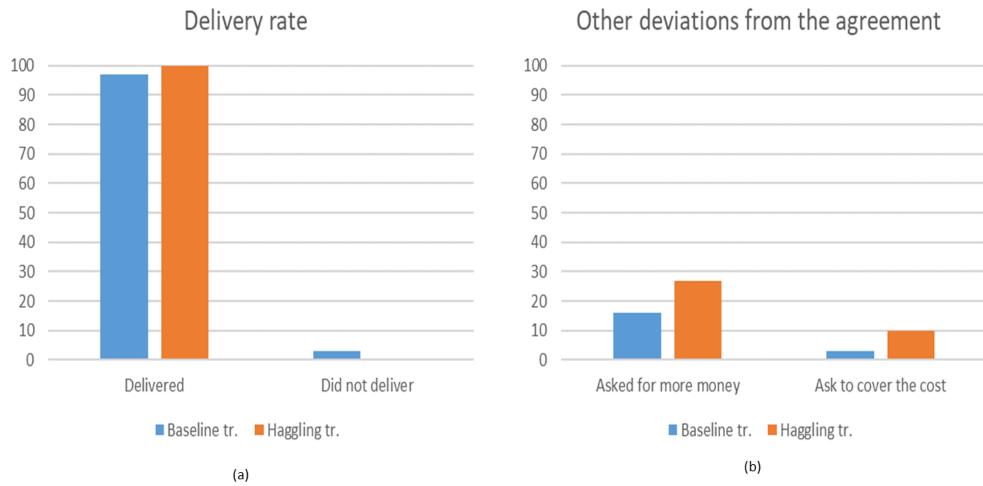


Figure 2 shows the delivery rate (panel a) and other breaches of the implicit agreement (panel b) between the RA-A and the cab driver.

Figure 2: Cab drivers' behavior



A.3 Two routes

Our main route follows Via Insurgentes, a major traffic artery in Mexico City. Via Insurgentes runs through some of the most affluent neighborhoods of the city.

Figure 3: Insurgentes route

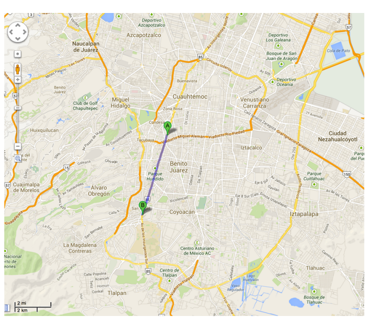
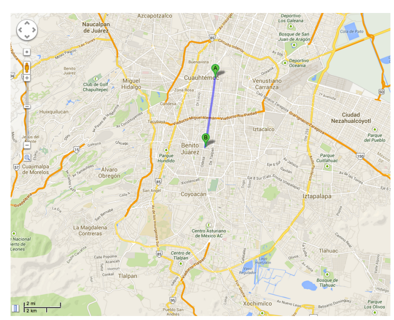


Figure 4: Downtown route



A.4 A survey

Would our findings be surprising also to MC locals? In a separate laboratory session conducted at ITAM, we invited 37 students, many of whom were MC residents, to shed light on that question. We asked these subjects to guess the outcome of the behavior in seven different scenarios involving plausible trust situations around MC; the fourth one was designed to be reminiscent of our experiment with the cab drivers:

1. You take a bus that is fully crammed with people. You manage to jump on in the back door. The only way for you to pay the fare is to send the money up front to the driver by asking passengers to pass it on up front. The bus fare is 5 pesos but you do not have any change so your only option is to pass 20 pesos. Will you get the change back?

2. You are at a street market (mercado sobre ruedas) looking for a gift for your friend. You come across a seller who is selling artisan tiles with custom writings on them. You would like to have one of those made with your friend's name on it. The seller wants the full payment of 100 pesos up front and promises to bring you the tile (same place same time) one week later. If you decide to go ahead and pay him, will you get your tile one week later?

3. You are at a football match and have to use a restroom. There is no assigned seating and you happen to have a good seat. If you leave your jacket on your seat as a place holder will it still be there after you've come back?

4. You need to pass a school project to your friend which is due the next day. The project contains sounds and video clips so you burn it on a CD. You cannot leave the house because you have to watch after your little brother. If you flag a cab on the street and pay the amount he asks ahead of time will the CD get to your friend?

5. You are in a bar around Centro Historico with a group of five friends celebrating your birthday. You feel generous and offer to pick up the tab for

the night. If you give the credit card to the bartender will the final tab at the end of the night be correct?

6. You go to the stadium to buy tickets for a football match that will take place tomorrow but they are sold out. A man outside the ticket office (a scalpel) offers to get you tickets at 30% discount. He does not have the tickets on him but has to walk over to his friend to get them. If you give him the money will he show up with the tickets?

7. You are going to a birthday party in La Condesa but can't find a spot to park your car. You have just come from a long road-trip and your car is full of personal belongings bags etc. in the back and front seat. If you leave the car at a Valet Parking will everything be there when you pick it up?

Responses were incentivized using the Houser & Xiao payment procedure according to which subjects get paid if their answers match the answer of another randomly selected participant in the room. This effectively creates a coordination game among the subjects. Our thinking was that if a subject thought it obvious that the correct answer would be yes, then they would attempt to coordinate on the corresponding equilibrium in the procedure. Overall, such coordination did not happen. In the fourth scenario, 43% of subjects guessed that the cab would not deliver; the rest guessed that the cab would deliver. Among those born in MC, 52% thought the cab would not deliver. This suggests that it is not common knowledge among MC locals that the correct answer should be yes.

A comparison across all seven scenarios provides some additional insights. We have listed them above according to how likely we found it (based on our intuition) to get a positive response, with scenario 1 appearing most likely to us. In the experiment, we scrambled the presentation order, and to allow some robustness check we ran two different orderings: Ord-#1 and Ord-#2. In Ord-#1 our cab scenario was listed second in the sequence and in Ord-#2 it was listed sixth. Ord-#1 was run with 18 subjects and Ord-#2 with

Table 1: Frequency of trusting responses

	Scenarios						
	1	2	3	4	5	6	7*
All data:							
Ord 1 (18)	13	13	6	4	1	10	10
Ord 2 (19)	17	12	8	8	2	10	11
Total	30	25	14	12	3	20	21
Total %	81	68	38	32	8	54	57
Mex. City natives:							
Ord 1 (14)	10	10	5	3	1	7	6
Ord 2 (13)	11	10	6	7	2	5	7
Total	21	20	11	10	3	12	13
Total %	78	74	41	37	11	44	48

Note: * denotes our cab scenario. In Ord 1 the cab scenario was 2nd in the sequence; in Ord 2 it was 6th. The number of observations for each ordering is in the parentheses.

19. The frequencies of trusting responses for all scenarios, orderings, and for full vs. restricted-to-MC-subjects sample are reported in Table 1. The frequencies of positive responses roughly align with our intuitions, with the cabs scenario taking a middling position.[15]

[15] Also: the responses are slightly more trusting in Ord #2 than Ord #1 and the observed patterns seem roughly consistent between the full and the restricted sample.

References

[1] Balafoutas L, Beck A, Kerchsbamer R, Sutter M (2013), "What drives taxi drivers? A field experiment on fraud in a market for credence goods," *Review of Economic Studies* 80, 876-891.

[2] Balafoutas L, Kerchsbamer R, Sutter M (2017), "Second-degree moral hazard in a real-world credence goods market," *Economic Journal* 127, 1-18.

[3] Bengtsson N, (2016), "Efficient informal trade: Theory and evidence from the Cape Town taxi market," *Journal of Development Economics* 115, 85-98.

[4] Berg J, Dickhaut J, McCabe K (1995), "Trust, reciprocity, and social history," *Games & Economic Behavior* 10, 122-142.

[5] Castillo M, Petrie R, Torero M, Vesterlund L (2013), "Gender differences in bargaining outcomes: A field experiment on discrimination," *Journal of Public Economics* 99, 35-482.

[6] Charness G, Dufwenberg M, (2006), "Promises and partnership," *Econometrica* 117, 817-869.

[7] Dufwenberg M, Kirchsteiger G (2004), "A theory of sequential reciprocity," *Games & Economic Behavior* 47, 268-98.

[8] Dufwenberg M, Servátka M, Vadovič R (2017), "Honesty and informal agreements," *Games & Economic Behavior* 102, 269-285.

[9] Fehr E, Gaechter S (2000), "Fairness and retaliation: The economics of reciprocity," *Journal of Economic Perspectives* 14, 159-181.

[10] Houser D, Xiao E (2011), "Classification of natural language messages using a coordination game," *Experimental Economics* 14, 1-14.

[11] Miettinen, T (2013), "Promises and conventions – an approach to pre-play agreements," *Games and Economic Behavior* 80, 68-84.

[12] Rabin, M (1993), "Incorporating fairness into game theory and economics," *American Economic Review* 83, 1281-1302.

[13] Sobel, J (2005), "Interdependent preferences and reciprocity," *Journal of Economic Literature* 43, 396-440.

[14] Vanberg C (2008), "Why do people keep their promises? An experimental test of two explanations," *Econometrica* 76, 1467-1480.