Behavior and Beliefs in Long-Distance Interactive Online Experiments between Moscow, Tomsk and Samara – First results

Alexis Belianin[a)], Heike Hennig-Schmidt[a,b)], Olga Kuznetsova[c)] Marina Ryzhkova[d)], Gari Walkowitz [e),] Gregory Chernov[a)]

[a)] International Laboratory for Experimental and Behavioural Economics, National Research University Higher School of Economics, Moscow, Russian Federation. [b)] Bonn University, Germany. [c)] Samara State University, Russian Federation. [d)] Tomsk State University, Russian Federation. [e)] Technical University Munich, Munich, Germany.

Draft version, not to be distributed

June 2019

Abstract

In this project, we report on a study on social preferences of individuals in three different parts of Russia (the cities of Moscow, Samara and Tomsk) by means of real economic interactions. Russia is a very large, which offers a unique chance to explore social preferences of people of the same nationality who most likely will never interact directly with each other in reality. Not much is known about interregional differences, especially those reflected in behavior. We analyze behavior by means of a laboratory experiment – the Ultimatum Game – that involves a monetary tradeoff between self-serving and other-regarding behavior. Our research questions relate to 1.) Heterogeneity: Do social preferences differ across geographically separated regions in Russia? 2.) Familiarity with own subject pool: Do social preferences and corresponding beliefs within a region match rather well? 3.) Alignment of actions and beliefs across subject pools: Do beliefs on matched counterparts' behavior from different Russian regions match that behavior? Our results show that differences across regions but also within regions exist. Particularly in Moscow, beliefs and decisions are not well calibrated.

1. Introduction

In this project, we report on a study on social preferences of individuals in three different parts of Russia (Moscow, Samara, Tomsk) by means of real economic interactions. Russia is a very large, yet relatively homogeneous country, which offers a unique chance to explore social preferences of people of the same nationality who most likely will never interact directly with each other in reality. Not much is known about interregional differences, especially those reflected in behavior. It may be, for instance, that idiosyncratic social norms existing within a region, related to fairness, reciprocity or cooperation differ from those in another region. Related existing measures show some cross-regional variance, but are limited to survey data. Long-distance behavioral experiments are a relatively new technique, which allows to elicit financially incentivized decisions between people and beliefs about each other in real time.

Our study will give new insights as only some studies on interregional differences in countries other than Russia exist (see section 2). On the country level, trust and cooperation has been shown to be associated with stronger economic performance, and to be stronger in countries that are less polarized, e.g., concerning ethnicity (Knack and Keefer 1997). Trust and cooperation can increase allocative efficiency by mitigating monitoring costs and contract enforcement problems (Herrmann et al. 2008). Hoewever, larger behavioral differences seem to exist within countries compared to differences between countries (Vieider et al., 2015, l'Haridon et al. 2017, Falk et al 2018). Finally, prosocial preferences of the Russians are relatively weak, related to observed behavior (Herrmann et al. 2008). Given these findings, sustainable development of Russian economy and society can be expected if cooperation and

trust among Russian citizens is stabilized and fostered by developing institutions supportive to these factors and by improving the understanding of fellow citizens' and aliens' behavior.

In our study, we specifically, analyze behavior by means of a laboratory experiment – the Ultimatum Game (Güth et al. 1982) – that involves a monetary tradeoff between self-serving and other-regarding behavior. In the Ultimatum Game, a Sender S is endowed by the experimenter with an amount of money that she can divide between herself and a Responder R. R can accept or reject the offer by stating the minimal offer he/she is willing to accept (MAO). In the former case both S and R get the amounts proposed, in the latter case S and R leave empty-handed. We also elicit incentivized belief data on the interacting partners' choices, both within and between regions. Further, we collected questionnaire data on personal characteristics and self-reported attitudes towards risk, trust, and life satisfaction.

The research questions we aim to answer in our study relate to 1.) Heterogeneity: Do behavioral measures of social preferences differ across geographically separated regions in Russia? 2.) Familiarity with own subject pool: Do social preferences and corresponding beliefs within a region match rather well? 3.) Alignment of actions and beliefs across subject pools: Do beliefs on matched counterparts' behavior from different Russian regions match that behavior?

We contribute to the literature by the first controlled experimental investigation comparing behavior and beliefs in an interactive within-country setting – also in Russia. We run the first within-Russia experiment with more than two subject pool and link our data to survey data as comprehensive surveys or field studies are missing. Further, we collect questionnaire data on personal characteristics and self-reported attitudes towards risk, trust, and life satisfaction. Our paper is substantially different from other studies in that we analyze actual within-country interactions. We use a controlled laboratory experiment run interactively and online via internet, in real time. We use student participants to have comparable subject pools

2. Related Literature

Behavioral differences might be expected because local norms can differ due to people acquiring ideas, beliefs and preferences from observation and interaction with other members of their own social group. Learning from peers can lead to stable social norms because people socially learn what is undesirable or even gets punished. These arguments have been put forward and have been supported in between-country studies (e.g. Bornhorst et al. 2010, Boyd and Richerson 1985, Richerson and Boyd 2005; Henrich et al., 2001; Boyd et al., 2003, Falk et al. 2018; Herrmann et al., 2008, Gächter and Herrmann 2011, Gäcjter et al. 2005; Goerg et al. 2016, Lönnqvist et al. 2015, Richerson et al. 2016, Cohn et al. 2019, Romano et al. 2017), in within-country analyses (e.g., Chmura et al. 2016; Cassar et al. 2014; Chua et al. 2019; Gallier et al. forthcoming; Michailidou and Rotondi, 2019; Kranton and Sanders, 2017, Liu and Zuo, 2019, Zhang 2015, 2018) and in within-city contexts (e.g., Falk and Zehnder 2013, Bigoni et al. 2018; Bogliacino et al. 2018; Rubin and Karaja 2018, Lei and Vesely 2010). When interactions are scarce, mistakes and misinterpretations are highly possible (Goerg et al., 2016), especially in such a large and diverse country as Russia. Given the findings in the literature our results are not unlikely to be supported for such a large and diverse country like Russia.

Not many experiments exist that use the technique of interactive simultaneous online long-distance behavioral experiments. Those that do exist are, e.g., Belianin and Novarese (2005): Inter-country Trust game, Russians and Italiens, two laboratories, 24 subjects; Goerg et al. (2016): Inter-country Trust game, Germans, Israelis and Palestinians, three laboratories, 400 subjects; Weimann et al. (2019): Public good game with large-groups (up to 100 subjects simultaneously), four laboratories in Germany, 5,160 subjects; Grimalda et al. (2018): Inter-country Collective risk social dilemma, Germans and Russians, four laboratories, two each in Germany and Russia, 784 subjects. In all of these large distance online

experiments at least one of the authors were involved (Belianin, Hennig-Schmidt, Ryzhkova and/or Walkowitz).

3. Regional differences from survey measures between the three locations

We chose Moscow, Tomsk and Samara, as these cities are rather far apart from each other: The distances are Moscow/Tomsk: 2,870 km, Moscow/Samara: 850, and Samara/Tomsk: 2,230 km (see Figure 1). Moreover, they are different in geographical locations, regional per-capita income, and number of inhabitants (see Table 1) that might give rise to behavioral differences and expectations. In particular, we looked at self-reported survey measures on trust, general and local solidarity as well as on united actions that are asked in the georating survey. These measures can be seen as background motivations relevant in the Ultimatum Game our participants are confronted with.

Table 1: Characteristics of the locations of our experiment

City (C)	Region (R)	Inhabitants (C) (01/01/2018)	Per-capita Income (R) (31/12/2017)
Moscow	Central region	12.51 Mio	62,532 RUR
Samara	Volga region	1.16 Mio	25,188 RuR
Tomsk	Western Siberia	0.57 Mio	24,457 RuR

Source: *https://ru-stat.com/*

Figure 1: Geographical locations of the three subject pools.

The georating survey data (2012) on general trust[1] shows that trust is relatively low in Russia (on average, about 19.3%), see Figure 3, and that trust differs substantially across Russian regions (see Figure 2).

[1] Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people? 1: Most people can be trusted. 2: Need to be very careful.

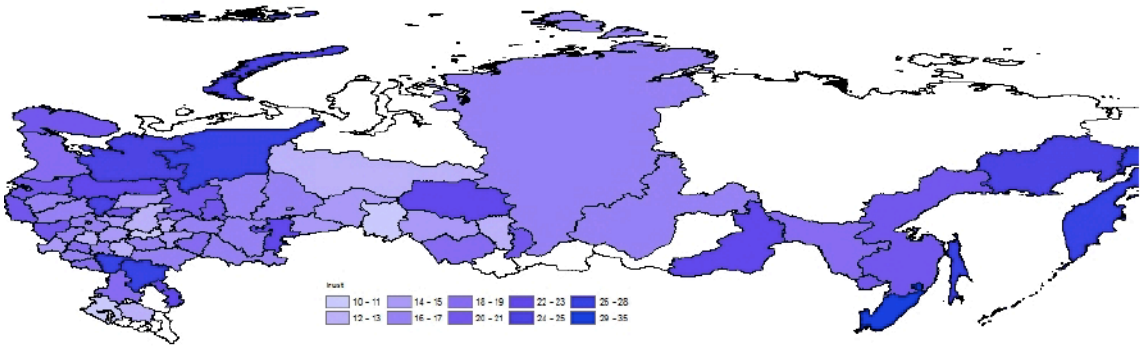

Figure 2: General trust across Russian regions (Source: Natkhov 2018)

Looking at the locations where we have run our experiment, we found that General trust is low in all three cities (see Figure 3) but it is significantly higher in Tomsk than in Moscow and Samara (p=0.0031 and p=0.0105, respectively, Chi2)[2]. We did not find a significant difference between Moscow and Samara (p=0.6733, Chi2).

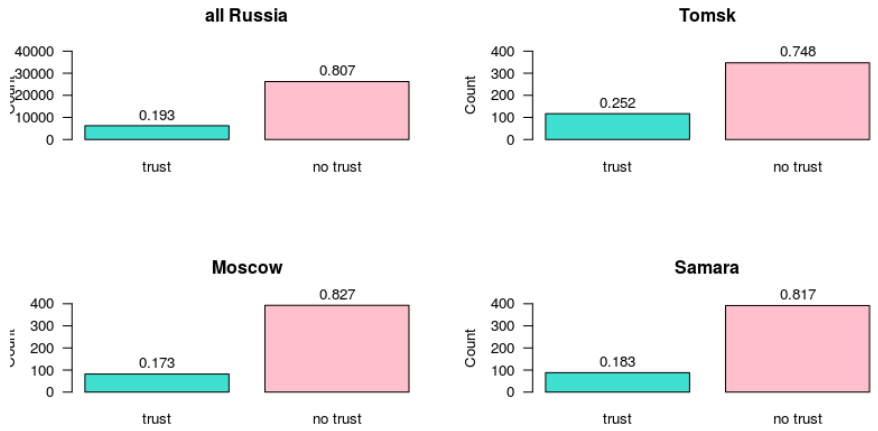

Figure 3: General trust stated in Russia, Tomsk Moscow, and Samara.
(georation 2012, own computations)

We next analyze how survey participants evaluate Solidarity in General: Do you think there is more harmony, cohesion or disagreement among people in our country today?[3]. We find a similar picture like with trust. Solidarity in General is seen to be low for Russia as a whole (about 17.6%, see Figure 4). Significant differences exist between Tomsk and Moscow as well as between Tomsk and Samara ($p < 0.001$) and no significant difference between Moscow and Samara (p=0.6733, both Chi2).

[2] All tests are two-sided throughout the paper if not mentioned otherwise.
[3] Possible answers were: Yes, rather Yes, rather No, No.

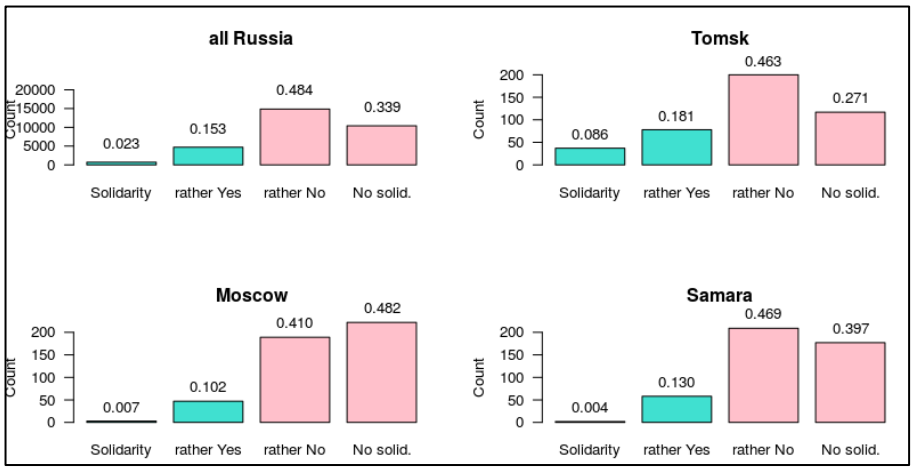

Figure 4: Solidarity in general stated in Russia, Tomsk Moscow, and Samara. (georation 2012, own computations)

The pattern is partially different when studying Local Solidarity: And if we talk about the people who surround you personally, among them today there is more agreement, solidarity or disagreement, disunity? 63. % of all Russian state Local Solidarity to be (rather) existent while these numbers are 76.0% in Tomsk, 71.6% in Moscow and 59.2% in Samara (see Figure 5). Local Solidarity is significantly higher in Tomsk and Moscow than in Samara ($p < 0.001$) and no significant difference between Moscow and Tomsk ($p=0.6733$, both Chi2).

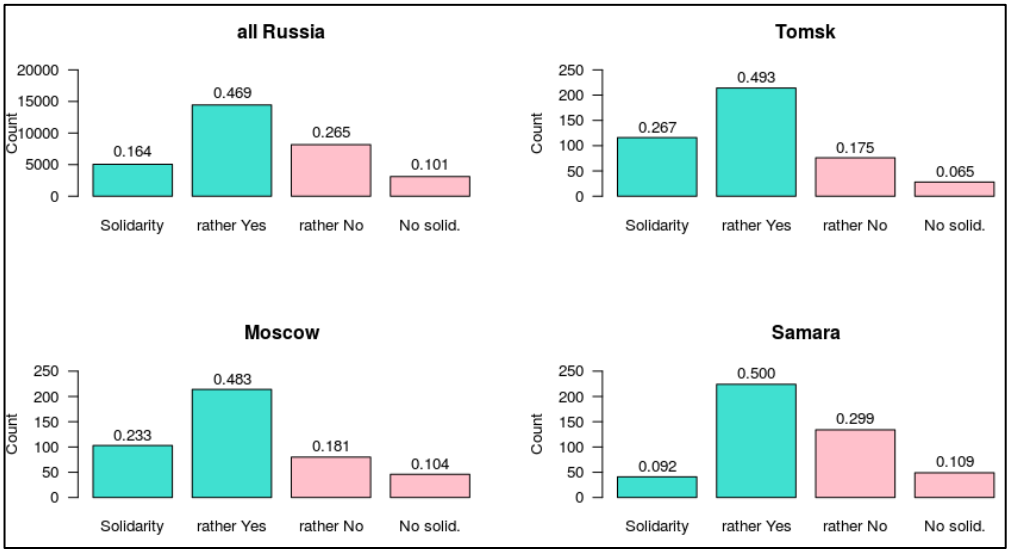

Figure 5: Local Solidarity stated in Russia, Tomsk Moscow, and Samara. (georation 2012, own computations)

Finally we look at how survey participants perceive themselves with regard to readiness for United Action: There are people who are ready to unite with other people for any joint actions, if their ideas and interests coincide. And there are people who are not ready to unite with others for any joint action, even if their ideas and interests coincide. To whom would you refer yourself - to the first or to the second?[4] In this respect, we found more Muscovites (76%) than people from Tomsk and Samara to be ready for joint

[4] Possible answers were: certainly to the first, rather to the first, more like second, of course the second

actions (69.3% and 65.8%, respectively, Figure 6). This difference is again significant ($p \leq 0.0335$, Chi2) with no difference to be found between the latter two cities (p=0.6733, Chi2).

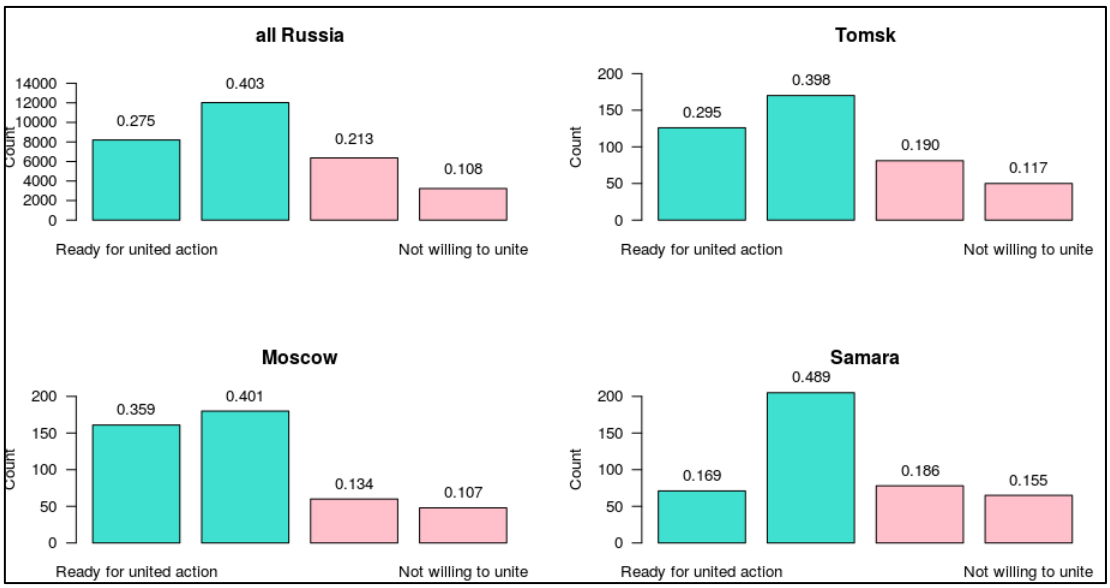

Figure 6: Readiness for United Action stated in Russia, Tomsk Moscow, and Samara. (georation 2012, own computations)

4. Research questions and conjectures

Survey evidence suggests heterogeneous levels of trust, local and general solidarity and readiness to united actions across the three cities. Our conjectures are as follows.

1. **Heterogeneity**: We expect the heterogeneity found in the survey data to be reflected in our behavioral data, i.e., our experimental subjects exhibit different levels of offers and MAOs across cities.
2. **Familiarity with own subject pool:** We conjecture standards of offers and MAOs to exist within the cities in that the distributions of Senders' offers/Responders' expectations on offers as well as Responders' MAOs and Senders' beliefs about MAOs within a subject pool do not differ significantly due to subjects being familiar with their own environment
3. **Alignment of actions and beliefs across subject pools:** We anticipate that beliefs on counterparts' behavior in other cities with whom a subject interacts do not match actual behavior in that cities if levels between cities differ since subjects are not familiar with norms in other cities.

5. Experimental design and procedure

5.1 Design

The workhorse we use to study our research questions is the Ultimatum game (UG) developed by Güth et. al (1982). A Sender S is endowed by the experimenter with an amount of money (the pie) that she can divide between herself and a Responder R. R can accept or reject the offer by stating the minimal offer he/she will accept (MAO). In the former case both S and R get the amounts proposed, in the latter case S and R leave empty-handed. We also asked for Senders' and Responders' expectations on counterparts decisions.

The sub-game perfect equilibrium is S keeping the whole pie, which R will accept. To exclude indifference between R accepting and rejecting a zero-offer, S will offer an amount equal to the smallest money unit which R will accept since he is better off than when rejecting. It has been shown in literally

thousands of UG experiments that the subgame perfect equilibrium has little predictive power. The probability of R's rejection is significantly higher if the offer becomes lower than 20 to 30% of the pie (Camerer 2003).

In our Ultimatum Game experiment, each Sender (Responder) plays three one-shot Ultimatum games with different parameters with the same Responder (Sender) in one session. Senders were named 'Participant 1' and Responders 'Participant 2' in order to not induce framing effects. Each S was endowed by the experimenter with a pie X she can divide between herself and R. R can accept or reject the offer. S's task is to decide on the share of X she is willing to offer to R. Simultaneously, R indicates the minimal offer he is willing to accept (Minimum Acceptable Offer, MAO; see Knez and Camerer 1995, Hennig-Schmidt et al. 2018). Each integer offer or MAO between 0 and X could be chosen. If S's offer exceeds R's MAO, X is divided according to the offer. Otherwise, both S and R receive no payoff from the game. See a sample decision screen for the Responder in Figure 7.

We also asked the Sender to state her first-order belief about the MAO of the Responder she is paired with. Likewise, we asked the Responder to indicate his first-order belief about the offer the paired Sender will make to him. Correctness of beliefs was incentivized by a reward according to the following Linear Absolute Value Rule (LAVR):

$$0.5 \times X \times \left(1 - \left| \frac{stated\ belief\ of\ R\ (S)\ on\ offer\ O\ (MAO)}{X} - \frac{factual\ offer\ O\ (MAO)\ of\ S\ (R)}{X} \right| \right).$$

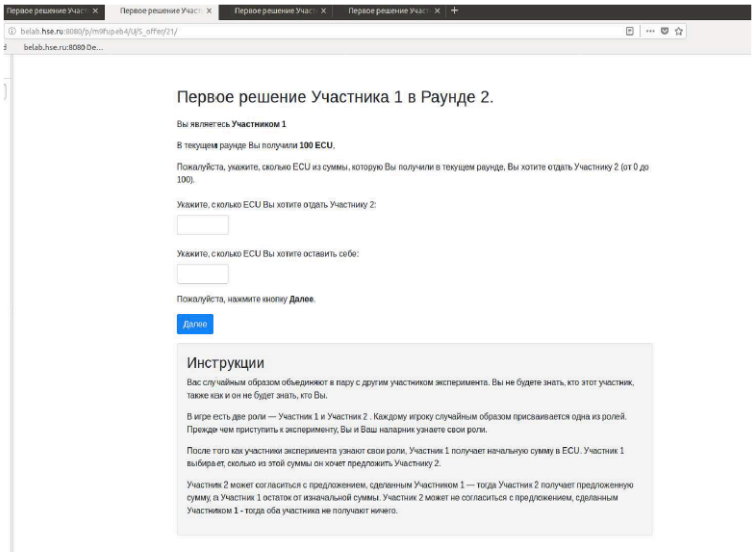

Figure 7: Sample Screen for Responders

The UG was played in a between-subject 3x3 design. In within-city sessions, participants at a location were randomly divided into two groups: Senders and Responders. In inter-city sessions, all participants in each of the two locations were randomly assigned to be either Senders or Responders.

Each Sender in each location (Moscow, Samara, Tomsk) made three consecutive decisions by interacting three times (i.e. in three rounds) with the same Responder from his own city or from one of the other two cities (see Figure 8). The same holds for Responders. In each of the rounds, the pie X increases from 60 Experimental Currency Units (ECU) in Round 1 over 100 ECU in Round 2 to 120 ECU in Round 3, see Table 2.

Table 2: Pie sizes in the Ultimatum Game

Round №	Pie size X in ECU
1.	60
2.	100
3.	120

One of the three rounds was randomly selected for payment, i.e. the payoff from the game plus the reward concerning the correctness of the belief. In addition, each subject was paid a show-up fee of 150 RuR. The exchange rate from ECU into RuR was 1: 5 in Moscow, i.e., the minimal pie size of 60 ECU = 300 RuR (5 euro), max pie size of 120 ECU = 720 RuR (10 euro). To account for different purchasing power in the respective cities, the exchange rate was 1: 3.5 in Tomsk and Samara (X = 60 ECU = 210 RuR, X = 120 ECU = 420 RuR).,

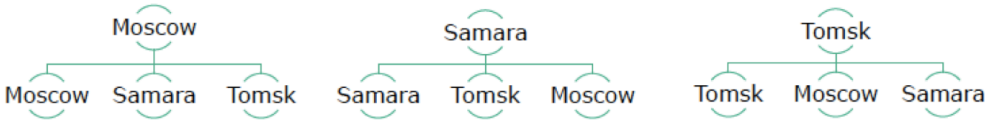

Figure 8: Matching of participants over locations

Feedback on payoffs and rewards was given only after the end of the last round. After having finished the experiment, subjects were asked to answer a questionnaire on personal characteristics and self-reported attitudes towards risk, trust, life satisfaction and freedom of choice.

5.2 Procedure

302 subjects participated in our UG experiment in total: 151 Senders and Responders each. These were 102 students from Moscow, 95 from Samara and 105 from Tomsk (Table 3). The participating laboratories are HSE Laboratory for Experimental and Behavioural Economics, Moscow, the Laboratory at the Department of Mathematical Methods in Economics of Samara State University and TUSUR Laboratory of Experimental Economics, Tomsk State University. We base our analysis on 14 – 20 independent observations per matching. As each participant took three decisions our data comprise 906 decisions in total.

After showing up for the experiment, participants were randomly allocated to their workstations and signed consent forms. In between-city as well as in within-city sessions, subjects were interacting with each other at the same time via online Internet connections. Subjects read the instructions, which gave detailed explanations on the experiment and the procedure. The experimenter then gave a recap of the instructions. Any questions participants might have had were answered in private.

We took great care to ensure that subjects understood the game and the incentive structure for rewarding the correctness of the belief elicitation task. To this end, we did not proceed in the experimental protocol until all subjects had correctly answered all online comprehension questions. Personal identities of matched participants were never revealed, however. To ascertain the credibility of the other lab's existence, both audiences saw each other via Skype. Except for the different exchange rates, all features of the experimental design and procedure were common knowledge and apparently did not raise any questions.

Table 3: Number of subjects by treatment and city

Treatment	Responder / Sender	Moscow	Samara	Tomsk	Total # Subjects	# Indep. Obs.	# Decisions/ player type
MosMos	Moscow	34			34	17	51
MosSam		16	16		32	16	48
MosTom		18		18	36	18	54
SamMos	Samara	15	15		30	15	45
SamSam			28		28	14	42
SamTom			16	16	32	16	48
TomMos	Tomsk	19		19	38	19	57
TomSam			20	20	40	20	60
TomTom				32	32	16	48
Total		102	95	105	302	151	453

Notes: Table entries correspond to numbers of players in each city pair. Senders' city is always listed first. Mos: Moscow; Sam: Samara, Tom: Tomsk.

Before making their decisions, subjects were informed where the respective counterpart (Responder/Sender) was located – either in their fellow-city or in one of the foreign cities. Then, Senders made their offer choices and Responders stated their MAOs (see the sample Screen for R in Figure 7). In each round, S and R were also asked to state their beliefs about the counter part's action. Subjects did not receive any feedback about their payoffs and rewards before the end of the third round of the experiment. One of the rounds was determined at random by the computer and the result was reported to both S and R. Finally, we asked subjects to fill in a questionnaire on personal characteristics and self-reported attitudes towards risk, trust, life satisfaction, and freedom of choice. At the end of a session, subjects were paid in private. Sessions lasted 70 – 80 minutes on average.

All decisions were anonymous and each subject participated in one session only. The experiment was programmed in oTree (Chen e.a., 2016) by which data exchange and communication of payoff calculations were executed. Sessions were run between April 2018 and March 2019. In total, Responders earned on average 345 RuR in Moscow, 285 RuR in Samara and 326 RuR in Tomsk for their decision and the belief elicitation task. Similarly, the payoffs for Senders were 338 RuR in Moscow, 246 RuR in Samara and 301 RuR in Tomsk. In addition, all participants received a show-up fee of 150 RuR in all three locations.

6. Results

Our focus in the result section is first on Senders' average offers and on Responders' average MAOs. We then analyze Responders' average expectations on offers and Senders' average expectations on MAOs. We use abbreviations for the different treatments. Note that the Senders' city is always listed first and Mos: Moscow; Sam: Samara, Tom: Tomsk. For instance, when analyzing offers, MosSam means that Senders from Moscow make offers to Responders from Samara, while when studying MAOs, MosSam indicates that Responders from Samara state their Minimal Acceptance Level of offers from Samarian Senders.

6.1 Senders' offers and Responders' minimal acceptable offers

We start with a descriptive analysis on players' choices and expectations on the counter players' first-order expectations. Altogether, 302 subjects with different study backgrounds took part in our

experiment, 102 from Moscow, 95 from Samara and 105 from Tomsk. 55.9% were males and the average age is 20.6 years. Recall that each subject made three decisions, one each for pie size $X = 60$, 100 and 120 ECU. We, therefore, have a total of 453 individual decision as well as expectation data each from Senders and from Responders – Senders' offers and expectations on MAO, Responders' MAO and expectations on Senders' offers. To normalize pie sizes we divide offers, MAOs and expectations by pie sizes. The resulting shares of the pie are our variables of interest and we report all results with regard to this variable. In the following analyses, for each individual we calculate the average over his/her three individual decisions as well as expectations to have independent observations.

Senders' average offers

We next analyze Senders' offers. Table 4 shows that averaged over all treatments, the level of Senders' offers is rather high by offering the Equal Split on average (50.5%). Comparing within-subject pool offers, we see that Senders from Samara offer significantly higher amounts to Responders from their own subject pool than Senders from Moscow and Tomsk offer to their fellow Responders (MosMos: 44.0%; SamSam: 52.1%, TomTom: 49.7%, see Table 4 and Figure 9). In particular, the difference between SamSam and MosMos is significant ($p = 0.010$, Mann-Whitney U-test, MWU in the following). We find no significant difference between Tomsk and Samara as well as between Tomsk and Moscow.

Comparing between-subject pool offers, we see that Muscovite Senders offer the lowest amounts of all three subject pools (see Table 4 and Figure 9). This concerns in particular their fellow responders. Differences across subject pools are not significant, however. Finally, Tomsk Senders tend to offer foreign Responders more than they offer to fellow-Responders.

Table 4: Senders' offers and Responders' expectations on Senders' offers, share of pie size

Treatment	# indepen-dent Obs.	Senders' offers		Responders' expectations	
		Mean	SD	Mean	SD
MosMos	**17**	**0.4342**	**0.0946**	**0.4448**	**0.1332**
MosSam	16	0.4629	0.0768	0.4910	0.2176
MosTom	18	0.4696	0.0550	0. 4802	0. 0602
SamMos	15	0.5446	0.1502	0.4937	0.1595
SamSam	**14**	**0.5151**	**0.0994**	**0.4956**	**0. 0473**
SamTom	16	0.5135	0.1277	0.4578	0.0884
TomMos	19	0.5563	0.2712	0. 4411	0.1293
TomSam	20	0.5610	0.1928	0.4941	0.0575
TomTom	**16**	**0.4804**	**0.0862**	**0.4387**	**0.1289**
Total	151	0.5054	0.1505	0.4708	0.1226

Notes: This table shows Senders' offers and Responders' expectations on Senders' offers in share of pie size differentiated by treatment. Senders' city is always listed first. Mos: Moscow; Sam: Samara, Tom: Tomsk. SD: standard deviation. **bold**: within-city interaction

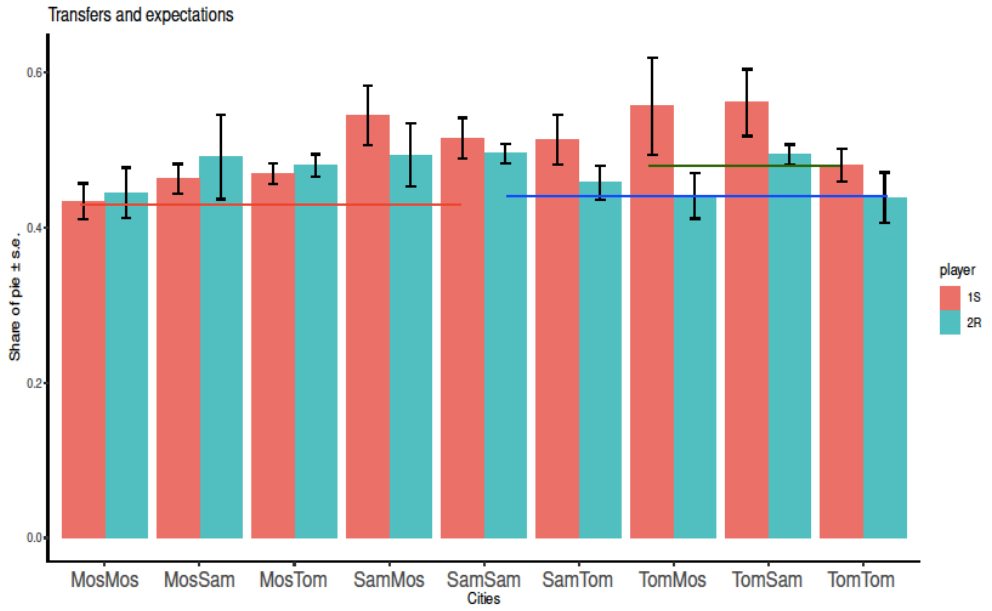

Figure 9: Senders' Offers and Responders' Expectations on Senders' Offers for averages over all pie sizes.

Notes: This Figure shows Senders' offers (red bars) and Responders expectations (green bars) in share of pie size averaged over all pie sizes and differentiated by treatment. Senders' city is always listed first. Mos: Moscow; Sam: Samara, Tom: Tomsk.

Responders' average MAOs

Our next point of interest is the amount up to which Responders are willing to reject an offer. We find more heterogeneity in MAOs than we have found in offers. Averaged over all treatments, the level of Responders' MAOs is rather high (36.2%, see Table 5) and definitely higher than the sub-game perfect equilibrium predicting that the Sender keeps (nearly) the whole pie, and the Responder accepts this division.

Table 5: Responders' Minimum Acceptable Offers (MAO) and Senders' expectations on Responders' MAOs, share of pie size

Treat-ment	# indep. Obs.	Responders' MAO		Senders' expectations	
		Mean	SD	Mean	SD
MosMos	**17**	**0.2664**	**0.1588**	**0.3873**	**0.1077**
MosSam	16	0.3432	0.0874	0.3302	0.1381
MosTom	18	0.3616	0.1491	0.3710	0.0929
SamMos	15	0.4110	0.2163	0.4531	0.2036
SamSam	**14**	**0.4470**	**0.1319**	**0.4496**	**0.0838**
SamTom	16	0.3769	0.1317	0.4394	0.0986
TomMos	19	0.3044	0.1728	0.3987	0.0889
TomSam	20	0.4046	0.1252	0.4251	0.1543
TomTom	**16**	**0.3641**	**0.0882**	**0.4209**	**0.1311**
Total	151	0.3621	0.1503	0.4072	0.1290

Notes: This Table shows Responders' Minimal Acceptance Levels (MAO) and Senders' expectations on Responders' MAOs in share of pie size differentiated by treatment. Senders' city is always listed first. Mos: Moscow; Sam: Samara; Tom: Tomsk. SD: standard deviation.

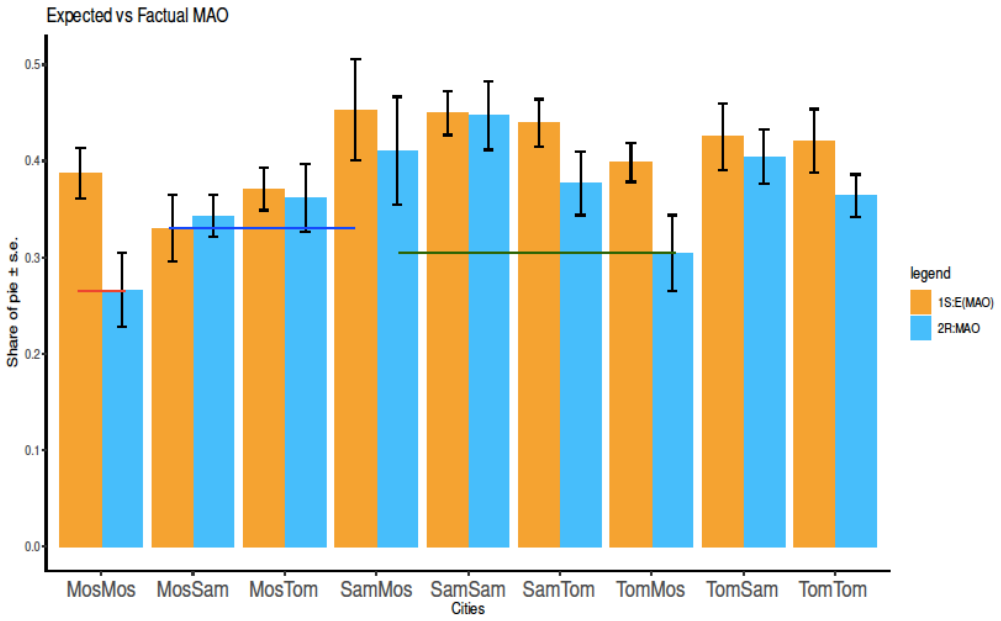

Figure 10: MAOs of Responders by treatments, and Expected Minimum Acceptable Offers (MAO) by Senders

Notes: This Figure shows Responders' MAO (blue bars) and Senders' expectations (orange bars) in share of pie size averaged over all pie sizes and differentiated by treatment. Senders' city is always listed first. Mos: Moscow; Sam: Samara, Tom: Tomsk.

In within-city interactions, Muscovite Responders state the lowest (26.6%) and Samara Responders the highest MAOs (44.7%) while Tomsk Responders' MAOs are in between (36.4%), see Table 5 and Figure 10. The within subject-pool difference in MAOs between Moscow and Samara is significant ($p = 0.0029$), those between Moscow and Tomsk as well as between Samara and Tomsk are weakly significantly different ($p \leq 0.0715$, all MWU-tests). We do not find significant differences in between-city interactions regarding MAOs.

We found Muscovite Responders' behavior to be noteworthy. Not only do they on average state the lowest MAOs of the within-city treatments. They also state the lowest MAO level compared to all other intercity interactions ($p \leq 0.0715$, MWU) except for MosSam and TomMos. Moreover, Muscovite Responders demand the highest of all stated average MAOs in inter-city interactions from Samara Senders (see Figure 10).

Result 1: **Different offer levels in Moscow and Samara**: Offers to fellow-Responders are significantly higher in Samara than they are in Moscow.
Result 2: **Heterogeneity in within-city MAO levels**: MAO levels differ between all three cities.
Result 3: Discrimination by Muscovite Responders: Muscovite Responders accept lower offers from their own and from Tomsk senders than from Samara.

Our above results partially support Conjecture 1 on Heterogeneity. With regard to offers, we found different levels between Moscow and Samara, while with regard to MAOs subject pools in all three cities differ significantly.

6.2 Responders expectations on senders' offers

The next step is analyzing Responders' expectations on Senders' offers, and in particular whether Responders correctly anticipate Senders offers in their own and in foreign subject pools.

As to within-city beliefs, Tomsk Responders state the lowest expectations (43.9%) while Samara Responders estimate the highest (49.6%) and Muscovite Responders are in between (44.5%), see Table 4 and Figure 9. Samara Responders expect to be offered nearly half of the pie from all three cities on average. The within subject-pool difference between Tomsk and Samara is significant ($p = 0.0300$). Those between Moscow and Tomsk as well as between Moscow and Samara are not ($p \geq 0.1091$, all MWU-tests).

Averaged over all treatments, Responders expect on average somewhat lower amounts than Senders actually offer (47.1 % vs. 50.5%). As to expectations in within-city treatments, Muscovite Responders slightly overestimate own Senders' offers, while Samarian and Tomsk Responders underestimate what their fellow-Senders' want to transfer to them. These differences are not significant, however ($p \geq 0.0993$, MWU[5]) meaning that Responders correctly anticipate Senders' offers from their own subject pool. The same holds for Responders' anticipation of Senders' offers from foreign cities, i.e., they actually estimate Senders offers rather correctly ($p \geq 0.0992$, MWU[6]), see Table 4 and Figure 9.

It is interesting to note that Responders from Moscow expect from own-city and from Tomsk Senders less (44.48%, and 44.11%) than Responders from any other subject pool do. On the other hand, they expect from Samarian Senders nearly half of the pie (49.56%). Both differences are not significant, though.

Result 4: **Well-calibrated expectation by Responders on own-subject pool Senders' offers**: Expectations are rather aligned within own subject pools. No differences in levels between cities exist.

Result 5: **Well-calibrated expectation by Responders on foreign-subject pool Senders' offers**: Expectations match Senders' offers in foreign cities quite well.

Result 4 and 5 support Conjecture 2 on Familiarity with own subject pool and Conjecture 3 on Alignment of actions and beliefs across subject pools. We anticipated that beliefs on counterparts' behavior in another city with whom a subject interacts do not match actual behavior in that city if levels between cities differ. In Moscow and Tomsk we do not find differences in within-city offer levels, and here expectations match offers quite well. Between Samara and Moscow we did find a difference in within-city offer levels. And in this treatment we find a rather weakly significant difference between offers and expectations.

6.3 Senders' expectations on Responders' MAOs

Finally, we analyze Senders' expectations on Responders MAOs, and in particular whether Senders' correctly anticipate their counterparts' minimal acceptable offers in their own and in foreign subject pools.

[5] The weakly significant difference concerns the Tomsk subject pools where offers are 4.17 percentage points higher than Responders expect.

[6] The weakly significant difference concerns treatment SamMos. In that respect it corresponds weakly to Result 1 showing a level effect between Moscow and Samara. Offers to fellow-Responders are significantly higher in Samara than they are in Moscow.

As to within-city beliefs, Muscovite Senders state the lowest expectations (38.7%), while Samara Senders estimate the highest (44.9%) and Tomsk Senders are in between (42.1%), see Table 5 and Figure 10. Again, people from Samara have the highest expectations, but this time with regard to MAOs, yet within subject-pool difference are not significant ($p \geq 0.1109$, MWU-tests). Interestingly, Moscow senders expect the lowest demands of Responders in within- and in interregional interactions compared to Samara and Tomsk

Averaged over all treatments, Senders anticipate higher amounts than Responders actually demand (36.2% vs. 40.7%). As to expectations in within-city treatments, Muscovite senders significantly and Tomsk Senders insignificantly overestimate own Responders' MAOs ($p = 0.0107$, and $p = 0.1086$, respectively, MWU), while Samarian senders are quite well calibrated ($p = 0.9444$, MWU). Calibration between Senders' anticipation of Responders' MAOs from foreign cities is rather good as Senders estimate Responders MAOs rather correctly ($p \geq 0.1738$, MWU).

Like with MAOs, we find more heterogeneity in Senders' expectations on MAOs across cities than for offers and expectations.

Result 6: Miscalibration between Moscow Senders' expectations and MAOs: Senders in Moscow significantly overestimate the amounts own Responders are demanding from them.

Result 7: Familiarity with own subject pool: In Samara and Tomsk Senders' anticipations and actual MAOs are well calibrated within their own subject pools.

Result 8: Alignment of actions and beliefs across subject pools: Except for Moscow, Senders' anticipations and actual MAOs are well aligned across subject pools.

Result 6 and 8 show that in Moscow senders' expectations and Responders MAOs are not well calibrated in within- and in between-city interactions. Apparently, Muscovite Senders overestimate Responders' willingness to accept low offers in their own subject pool. This result contradicts Conjecture 2 on Familiarity with own subject pool. This is different in Samara and Tomsk where expectations and MAOs are (rather) similar in within-city and in between-city interactions. Insofar, Conjectures 2 and 3 (Alignment of actions and beliefs across subject pools) are supported for Samara and for Tomsk.

References

Belianin and Novarese (2005): Trust, communication and equlibrium behaviour in public goods game: a cross-country experimental study, Working paper.

Bigoni, M., S. Bortolotti, and V. Rattini (2018): A Tale of Two Cities: An Experiment on Inequality and Preferences, Bucconi University, IZA DP No. 11758, http://ftp.iza.org/dp11758.pdf

Bogliacino, F., C. Gómez, F. Montealegre and G. Grimalda (2019): A city-wise experiment on trust, beliefs, and collective perceptions, paper presented at Centro de Investigaciones para el Desarrollo, Universidad Nacional de Colombia.

Bornhorst F., A. Ichino, O. Kirchkamp, K.H. Schlag, and E. Winter (2010). Similarities and differences when building trust: The role of cultures. *Experimental Economics,* 13, 260-283, https://link.springer.com/article/10.1007%2Fs10683-010-9240-x.

Boyd, R. and P.J. Richerson (1985): Culture and the Evolutionary Process, The University of Chicago Press, London, Chicago;

Boyd, R., Gintis H., Bowles S. and P.J. Richerson (2003): The evolution of altruistic punishment, Proc Natl Acad Sci USA, 3531 – 3535, doi 10.1073/pnas.0630443100

Cassar, A., d'Adda, G., Grosjean, P. Institutional quality, culture, and norms of cooperation: Evidence from behavioral fieldexperiments (2014) Journal of Law and Economics, 57 (3), pp. 821-863. http://dx.doi.org/10.1086/678331

Chen, D.L., Schonger, M., Wickens, C., (2016): oTree—An open-source platform for laboratory, online, and field experiments. Journal of Behavioral and Experimental Finance 9 (1), 88-97.

Chen, D.L., Schonger, M., Wickens, C., (2016): oTree—An open-source platform for laboratory, online, and field experiments. Journal of Behavioral and Experimental Finance 9 (1), 88-97.

Chmura, T., S. J. Goerg, and P. Weiss (2016), Natural groups and economic characteristics as driving forces of wage discrimination, *European Economic Review*, 90, 2016, 178-200, https://www.sciencedirect.com/science/article/pii/S0014292116300708

Chua, R.Y.J. K. G. Huang, M. Jin (2019): Mapping cultural tightness and its links to innovation, urbanization, and happiness across 31 provinces in China, Proceedings of the National Academy of Sciences 116 (14) 6720-6725; DOI: 10.1073/pnas.1815723116

Cohn, A., Michel André Maréchal, David Tannenbaum, Christian Lukas Zünd (2019): Civic honesty around the globe, Science, DOI: 10.1126/science.aau8712

Falk, A., Becker A., Dohmen T. , B. Enke, D. Huffman, and U. Sunde (2018): Global Evidence on Economic Preferences, Quarterly Journal of Economics, 133(4), 1645-1692, 2018

Falk, A., and C. Zehnder, (2013): A city-wide experiment on trust discrimination, *Journal of Public Economics*, 100, 2013, 15-27, https://www.sciencedirect.com/science/article/abs/pii/S0047272713000182

Gächter, S., Herrmann, B. and Thöni, C. 2005: Cross-cultural differences in norm enforcement. Behav. Brain Sci., 28, 822

Gächter, Simon, Benedikt Herrmann (2011): The limits of self-governance when cooperators get punished: Experimental evidence from urban and rural Russia, European Economic Review, Volume 55, Issue 2, 193-210,

Gallier, C., T. Goeschl, M. Kesternich, J. Lohse, C. Reif und D. Römer (fortcoming), Leveling Up? An Inter-Neighborhood Experiment on Parochialism and the Efficiency of Multi-level Public Goods Provision, *Journal of Economic Behavior and Organization* (http://ftp.zew.de/pub/zew-docs/dp/dp17012.pdf).

georating (2012): Public Opinion Foundation, Georating poll.

Goerg, S. J., H. Hennig-Schmidt, G. Walkowitz, E. Winter (2016): In Wrong Anticipation – Miscalibrated Beliefs between Germans, Israelis, and Palestinians. PLoS ONE 11(6): e0156998 (dx.doi.org/10.1371/journal.pone.0156998)

Grimalda , G., A. Belianin, H. Hennig-Schmidt, T. Requate, M. V. Ryzhkova (2018): Sanctions and convergence of cooperative behavior improve international collaboration – An interactive experimental study within and between Germany and Russia. Manuscript presented at the World Meeting of the Economic Science Association 2018, Berlin Germany.

Güth. W., R. Schmittberger, B. Schwarze. B. (1982): An Experimental Analysis of Ultimatum Bargaining. Journal of Economic Behavior and Organization. 3. 367-388.

Hennig-Schmidt, H., B. Irlenbusch, R. M. Rilke, and G. Walkowitz (2018): Asymmetric Outside Options in Ultimatum Bargaining: A Systematic Analysis, International Journal of Game Theory, 47, 301–329, doi.org/10.1007/s00182-017-0588-4.

Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, R. McElreath, (2001): In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies. American Economic Review, Papers and Proceedings, 91, 73-78.

Herrmann, B., Ch. Thöni, S. Gächter (2008): Anti-social punishment across societies, Science 319, 1362-1367.

Knack S, Keefer P. 1997. Does social capital have an economic payoff? A cross-country investigation. Quarterly Journal of Economics, 112:1252–88

Knez MJ, Camerer CF (1995): Outside options and social comparison in three-player ultimatum game experiments. Games Econ Behav 10(1):65–94

Kranton, R.E., and S.G. Sanders (2017): Groupy versus non-groupy social preferences: personality, region, and political party. *American Economic Review,* 107, 65–69, https://www.aeaweb.org/articles?id=10.1257/aer.p20171096., https://sites.duke.edu/rachelkranton/files/2016/09/Groupy-vs.-NonGroupy-Social-Preferences-Personality-Region-and-Political-Party.pdf

l'Haridon, Olivier, Vieider, Ferdinand M., Aycinena, Diego, Bandur, Agustinus, Belianin, Alexis, Cingl, Lubomír, Kothiyal, Amit and Martinsson, Peter (2018): Off the Charts: Massive Unexplained Heterogeneity in a Global Study of Ambiguity Attitudes,The Review of Economics and Statistics, 100, 4, 664-677, 10.1162/rest_a_00724

Lei, V., and F. Vesely (2010). In-Group versus Out-Group Trust: The Impact of Income Inequality. *Southern Economic Journal, 76*(4), 1049-1063. https://onlinelibrary.wiley.com/doi/pdf/10.4284/sej.2010.76.4.1049,

Liu, E. M., and S. X. Zuo (2019): Measuring the impact of interaction between children of a matrilineal and a patriarchal culture on gender differences in risk aversion, Proceedings of the National Academy of Sciences, 116 (14) 6713-6719; DOI: 10.1073/pnas.1808336116

Lönnqvist, J.-E., H. Hennig-Schmidt, G. Walkowitz (2015): Ethnicity- and Sex-Based Discrimination and the Maintenance of Self-Esteem, *PLoS ONE* 10(5): e0124622 (doi.org/10.1371/journal.pone.0124622).

Michailidou, G. and V. Rotondi (2019): I'd Lie for You, European Economic Review, 118, 181-192, https://www.sciencedirect.com/science/article/pii/S0014292119301072.

Natkhov, T.V. (2018): Education and Social Capital in Russia – Empirical Evidence. Paper presented at International Workshop "From values to behaviour: interdisciplinary studies of regional differences in Russia", Moscow, 23 October 2018, NRU Higher School of Economics, Moscow, Russian Federation.

Richerson et al (2016): Cultural group selection plays an essential role in explaining human cooperation: A sketch of the evidence, Behavioral and Brain Sciences (2016), Page 1 of 68, doi:10.1017/S0140525X1400106X, e30

Richerson, P.J. and R. Boyd (2005): Not by the genes alone – How culture transforms human evolution. The University of Chicago Press, London, Chicago.

Romano, A., Balliet, D., Yamagishi, T., & Liu, J. H. (2017). Parochial trust and cooperation across 17 societies. Proceedings of the National Academy of Sciences, 114(48), 201712921. http://doi.org/10.1073/pnas.1712921114

Rubin, J. and Karaja, E. (2018): The Cultural Transmission of Trust Norms: Evidence from a Lab in the Field on a Natural Experiment, Working paper, Chapman University, https://www.chapman.edu/research/institutes-and-centers/economic-science-institute/_files/WorkingPapers/karaja-rubin-2017.pdf

Vieider, F. M., Lefebvre, M. , Bouchouicha, R. , Chmura, T. , Hakimov, R. , Krawczyk, M. and Martinsson, P. (2015), Common components of risk and uncertainty attitudes across contexts and domains: evidence from 30 countries. *Journal of the European Economic Association*, 13: 421-452. doi:10.1111/jeea.12102,

Weimann J., J. Brosig-Koch, T. Heinrich, H. Hennig-Schmidt, C. Keser. (2109): The Logic of Collective Action Revisited, *European Economic Review*, 118, 348-363, https://doi.org/10.1016/j.euroecorev.2019.05.019

Zhang, N. (2015). Changing a 'culture' of corruption: Evidence from an economic experiment in Italy. *Rationality and Society*, 2015, Vol. 27(4) 387–413 DOI: https://doi.org/10.1017/XPS.2017.26

Zhang, N. (2018). Institutions, Norms, and Accountability: A Corruption Experiment with Northern and Southern Italians. *Journal of Experimental Political Science, 5*(1), 11-25. doi:10.1017/XPS.2017.26