ESSAYS ON SOCIAL COHESION

A Dissertation
Presented by the Faculty of the Department of Economics and Statistics

of

University of Naples Federico II

in Candidacy for the Degree of

Doctor of Philosophy

by

Federica Ambrosio

May 2025

Abstract

Essays on social cohesion

Federica Ambrosio

2025

The dissertation includes three chapters on applied microeconomics and development economics, with a focus on the relationship between social cohesion and growth opportunities in developing countries. The first chapter reports new evidence about the relationship between perceived economic status and social trust and shows that biases in perceived economic status are relevant in studying economic interactions and policy analysis. The second chapter explores the relationship between social cohesion and informal risk-sharing engagements. By comparing theoretical predictions on the incentives to share risk with empirical laboratory data collected in Colombia, the chapter highlights the role of social cohesion in improving the efficiency of risk-sharing. Lastly, the third chapter include empirical tests on the causal relationship between agricultural production and social cohesion in Colombian municipalities, showing that a higher use of labour-intensive agriculture can increase social cohesion in rural areas.

Contents

List of Figures

List of Tables

Introduction

The dissertation covers topics related to social cohesion and development economics. The three chapters are inspired by the psychology literature on altruism and social identity (e.g. Tajfel (1970), Batson (1990), Batson (2010)), the economic literature on cooperation in informal settings (e.g. Berg, Dickhaut, and McCabe (1995), Foster and Rosenzweig (2001), Townsend (1995)), and the economic literature on networks and risk preferences (e.g. Bramoulle and Kranton (2007), Demuynck, Herings, Saulle, and Seel (2019), Jackson, Rodriguez-Barraquer, and Tan (2012)).

The objective is to explore the role of social cohesion in creating growth opportunities in developing countries (e.g. Alan, Baysan, Gumren, and Kubilay (2021), Benabou and Tirole (2006), Berg et al. (1995)). In particular, the three essays focus on the relationship between social cohesion and individual economic preferences, deepening the practical implications of this relationship in the context of risk-sharing (e.g. Townsend (1995), Bramoulle and Kranton (2007), Fafchamps and Gubert (2007)), and exploring the role that agricultural practices historically had in shaping social cohesion in emerging markets (e.g. Rodriguez-Giraldez and Vicente-Ramos (2020), Vollrath (2011), Fiszbein, Jung, and Vollrath (2022)).

In Chapter 1, the dissertation begins with reporting empirical evidence on the relationship between perceived economic status and general levels of trust

in a society, using survey data from the 2016 wave of the Life in Transition Survey (World Bank and EBRD). In the economic literature, trust is considered a relevant factor in the development of countries lacking strong formal institutions (Dearmon and Grier (2009)) and the study of its determinants can have relevant policy implications. Contrarily to the standard economic literature on trust, which highlights actual economic success (Alesina and La Ferrara (2002)) and social status (Lount and Pettit (2012)) as predictors of trust, this work addresses the role of perceived economic status and its biases. The analysis shows that perceived economic status correlates more strongly with trust than actual economic status. Specifically, a negative perception bias correlates with lower trust levels. Moreover, it is shown that distortions in perceived economic success are strongly correlated with perceived intergenerational mobility, fairness of institutions and network ties.

In Chapter 2, the dissertation follows with a theoretical and empirical analysis of the effects of social cohesion on informal risk-sharing contracts, using experimental data collected in Colombia. The economic literature agrees with informal risk-sharing being a fundamental risk coping mechanism in developing countries, which are often characterised by a lack of formal insurance mechanisms for the rural poor (e.g. Fafchamps and Gubert (2007), Townsend (1995)). However, empirically, informal risk-sharing is almost exclusively observed within networks of households and friends. Chapter 2 aims to understand the incentives underlying this practice. Two mechanisms are analysed: (1) reciprocal trust and (2) social costs at the network level. To test each mechanism, I use the information collected through a lab-in-the-field experiment run in Colombia. Thanks to the data available, I can proxy for the level of collective and individual trust. These data, together with information on network ties, are then used to estimate the individual likelihood of forming a

risk-sharing group. The analysis shows that both social costs and trust are relevant in explaining risk-sharing arrangements. However, trust appears as a sufficient condition to join risk-sharing groups in absence of social costs. This indicates that it is possible to improve risk-sharing efficiency by creating conditions that increase overall trust in society, encouraging people to engage in risk-sharing activities outside their networks of family and friends.

To conclude, Chapter 3 introduces a geographical study of the municipalities analysed in Chapter 2, using Colombian census data to explore the possible determinants of social cohesion and collective preferences at the municipality level. In particular, I test the hypothesis that the prevalence of labour-intensive agriculture contributed to the development of social cohesion due to the repeated interactions between stable and dedicated groups (Vollrath (2011)). Compared to the existing literature, this study relies on a direct measure of social cohesion collected through a lab-in-the-field experiment. To avoid endogeneity issues, an instrumental variable strategy is implemented, instrumenting the presence of labour-intensive agriculture with historical precipitation data, under the assumption that the level of precipitations directly affects specific crop suitability (Fischer, van Velthuizen, Shah, and Nachtergaele (2002)). The empirical analysis shows that areas with a prevalence of labour-intensive agriculture exhibit higher levels of social cohesion, confirming the influence of agricultural practices on collective preferences (Fiszbein et al. (2022)).

Chapter 1

Trust, economic status and perception bias

1.1 Introduction

Trust is widely analysed in the economic literature due to its implications in predicting economic behaviours (e.g. Alesina and La Ferrara (2002), Glaeser, Laibson, Scheinkman, and Soutter (2000), Sapienza, Toldra-Simats, and Zingales (2013)). Most importantly, as Arrow (1974) points out, trust is a lubricant for social systems and boosts economic interactions among individuals.

In development economics, as Dearmon and Grier (2009) highlight in their study on the links between trust and economic growth, trust directly influences investments in physical and human capital. Hence, the study of the factors influencing trust has a relevant role in understanding the foundations of economic behaviour and countries' growth opportunities (Berg et al. (1995), Siegrist (2021)).

Contrarily to the standard economic literature on trust, which stresses how actual economic success (Alesina and La Ferrara (2002)) and social status

(Lount and Pettit (2012)) are among the most relevant predictors of trust, this chapter focuses on the role of perceived economic status and adds new insights on the relationship between perception biases and trust.

Indeed, as Graafland and Lous (2019) notice, economic success depends on one's actual position on the income ladder and its individual perception. However, the two might not always coincide. In this respect, the economic literature on trust has often ignored the role of distortions in perceived economic success.

On the other hand, perception biases have been widely studied in the psychological and medical literature, emphasizing how, compared to the measures of objective socioeconomic status (SES), subjective SES is a more sensitive predictor of psychological and physiological functioning (Adler, Castellazzo, Epel, and Ickovics (2000), Marmot (2006)). Inspired by these findings, this chapter aims at studying the effects of misperception of economic status on the individuals' likelihood to trust others.

The study exploits survey data collected jointly by the European Bank for Reconstruction and Development (EBRD) and the World Bank, administered in 2016 to a representative sample of households in countries transitioning from communism to market economy (Life in Transition Survey - LiTS)[1]. Transitioning countries, due to the structural changes happening in the economy, are more likely to have individuals with temporarily misaligned perceptions of income distribution in the country, leaving more space for the study of its variations.

The analysis compares measures of actual economic success to survey responses about subjective economic status, the difference between the two is

[1]European Bank for Reconstruction and Development, World Bank. Life in Transition Survey (LiTS) 2016, https://www.ebrd.com/what-we-do/economic-research-and-data/data/lits.html, accessed Sept 13, 2023.

then used as a proxy for perception bias. Perception bias is here intended, following Dasgupta and Saha (2022), as a misperceived self-efficacy, which the authors highlight as a strong determinant of intergenerational transmission of poverty. Indeed, results identify education and gender among the individual characteristics that are most correlated with the perception bias (see Appendix 1.6). These characteristics, as the literature indicates, are strongly correlated to one's self-perception (Exley and Nielsen (2024), Filippin and Paccagnella (2012)).

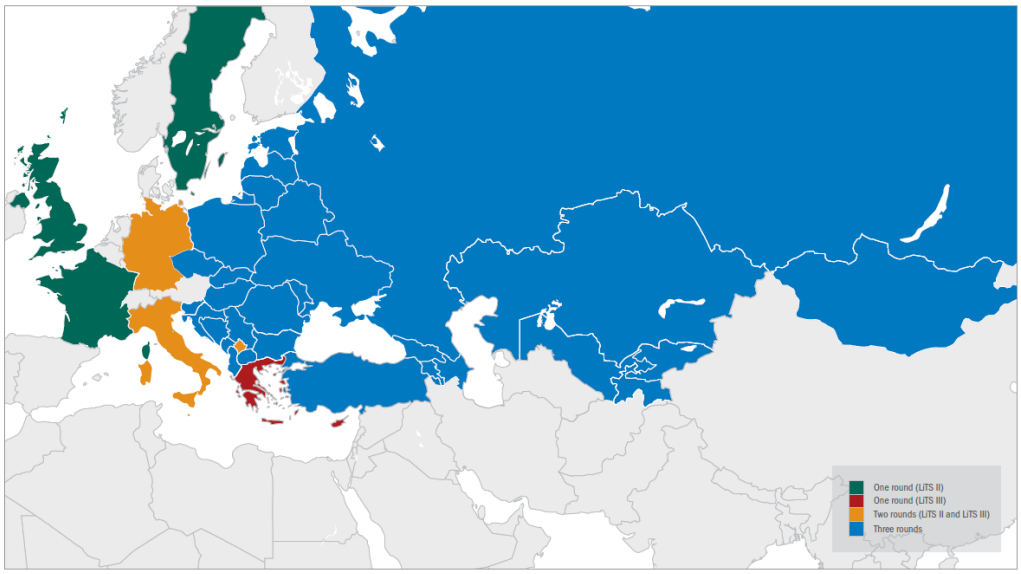
Overall, this chapter shows that subjective economic success correlates more strongly with variations in trust than actual economic success. In particular, subjective economic status positively correlates with a variation of around 6% in trust, compared to a 3% variation correlated with actual economic status, and the R^2 of the regression on trust increases by around 2 percentage points when substituting perceived economic status to actual economic status. Lastly, this chapter sheds light on the relevance of perceived intergenerational mobility, fairness of institutions, and network ties in understanding variations in the perception bias.

Results have important policy implications, suggesting that countries with corrupted institutions and weak network ties are more likely to fall into a vicious cycle of pessimism about personal position on the economic ladder, low levels of trust, and fewer economic relationships. In this respect, it appears fundamental in policy analysis to account for individuals' satisfaction with their social environment and how they perceive their role in the society.

The rest of the chapter includes a section describing the measures collected for the empirical analysis (Section 1.2), followed by a section on the empirical strategy (Section 1.3) and a section on the results (Section 1.4).

1.2 Measures of trust and economic success

Data are collected through the LiTS survey, administered by the EBRD and the World Bank to households in countries transitioning from socialism to market economies. The household survey is combined with an attitudinal survey that helps exploring the links between economic conditions and individual characteristics. The survey was administered in 2006, 2010, and 2016 to different households, randomly selected according to countries' representativeness. Figure 1.1 shows the countries' involved in the three waves. This study focuses on the 2016 wave, as it includes detailed information on household income, which was not collected in the other waves. The 2016 survey (LiTS III) interviewed 51,000 households in 34 countries.



Source: LiTS I (2006), LiTS II (2010) and LiTS III (2016).

Figure 1.1: Countries included in the survey

The information used in the empirical analysis refers to trust and income ranking measures. The main measure of trust used is the answer to the following question: *"Generally speaking, do you think that most people can be trusted or that you can't be too careful in dealing with people?"*.

This question is widely used in surveys, like World Value Survey and General Social Survey, to capture trust and trustworthiness (e.g. Glaeser et al. (2000), Lazzarini, Madalozzo, Artes, and Siqueira (2004)). However, as Glaeser et al. (2000) suggest, the responses can be subject to different interpretations. Indeed, interviewees' answers may depend on how they interpret "most people", what group of individuals they have in mind when answering the question, and their definition of "trust." For this reason, the following other questions are considered in the analysis to check for the results validity.

- *To what extent do you trust people from the following groups?*
  - family living with you
  - your neighbourhood
  - people you meet for the first time
  - foreigners

- *Suppose you lost your (purse/wallet) containing your address details, and it was found in the street by someone living in this neighbourhood. How likely would it be returned to you with nothing missing?*

These two additional questions leave less space for subjective interpretations since they ask about specific groups of people and situations involving trust. The first question, indeed, defines the groups of reference, while the second question gives a precise definition of trust, which mimics the "trust game" designed by Berg et al. (1995).

Figure 1.2 reports the distribution of the answers for all the measures of trust used. It is worth noticing how the distribution of responses to the question *"Most people can be trusted"* and the question about the likelihood of a wallet being returned are similar, as is the distribution of trust towards newly met people and foreigners, suggesting a similar interpretation by respondents for these two groups of questions.

On the other hand, as one can expect, trust towards family members is concentrated on the "complete trust" side. A similar skewness is registered in the answers regarding trust towards neighbours, with a higher density for the answer "some trust" than "complete trust".

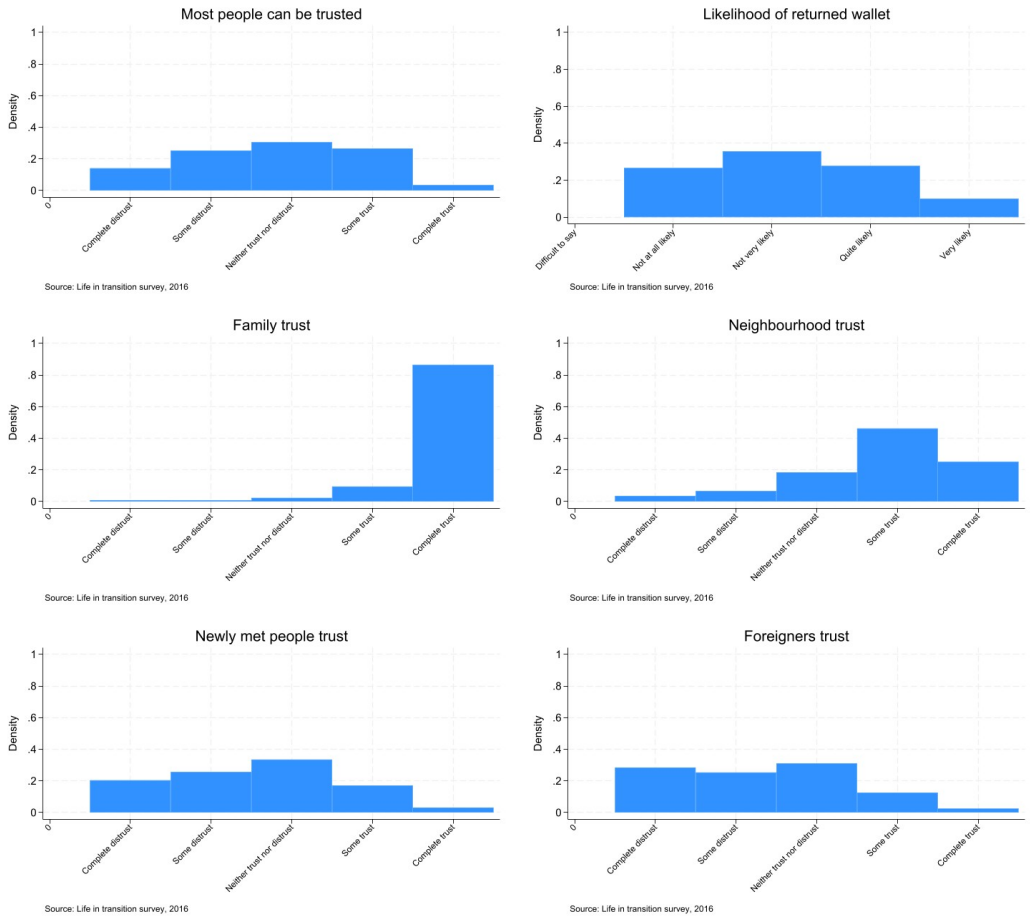


Figure 1.2: Measures of trust

Notes: *"Most people can be trusted"* shows the distribution of responses to the question *"Generally speaking, do you think that most people can be trusted or that you can't be too careful in dealing with people?"*."*Likelihood of returned wallet"* shows the distribution of responses to the question *"Suppose you lost your (purse/wallet) containing your address details, and it was found in the street by someone living in this neighbourhood. How likely would it be returned to you with nothing missing?"*."*Family trust"* shows the distribution of responses to the question *"To what extent do you trust the family living with you?"*."*Neighbourhood trust"* shows the distribution of responses to the question *"To what extent do you trust people from your neighbourhood?"*."*Newly met people trust"* shows the distribution of responses to the question *"To what extent do you trust people you meet for the first time?"*."*Foreigners trust"* shows the distribution of responses to the question *"To what extent do you trust foreigners?"*.

In the empirical strategy in Section 3.3, all answers are categorized in the

following dichotomic format:

$$trust = \begin{cases} 0 & \text{some/complete distrust} \\ 1 & \text{some/complete trust} \end{cases}$$

$$wallet = \begin{cases} 0 & \text{not at all/not very likely} \\ 1 & \text{quite/very likely} \end{cases}$$

In order to measure the actual and subjective income rank, the following variables are used. The first variable, used to measure the subjective income rank, is the answer to the following question: *"Imagine a ten-step ladder where on the bottom, the first step, stand the poorest 10% people in our country, and on the highest step, the tenth, stand the richest 10% people in your country. On which step of the ten is your household today?"*. The question is based on the subjective socioeconomic scale first introduced by Adler et al. (2000), to measure interviewees' perceived economic status.

On the other hand, to measure the actual income rank, the income decile distribution is calculated in each country based on the weighted data on households' registered income, and each household is assigned to the decile of belonging. To compare the measures of actual and subjective income rank, both variables are standardised. Lastly, the difference between the two variables is calculated to measure the *perception bias* of each household in the sample regarding their income status.

$$\text{perception bias} = \text{subjective income rank} - \text{actual income rank} \tag{1.1}$$

For interpretative reasons, the *perception bias* indicator is reported as follows in the empirical analysis (Section 3.3):

$$\text{perception bias} = \begin{cases} 1 & \text{negative (i.e. subjective income lower than actual income)} \\ 2 & \text{null (i.e. subjective income equal to actual income)} \\ 3 & \text{positive (i.e. subjective income higher than actual income)} \end{cases}$$

As shown in Figure 1.3, the correlation between average subjective income rank and actual income rank is positive but weak. The figure shows a broader variation in the distribution of subjective ranks compared to actual income across countries. This is even more evident when looking at the distribution within countries, which reports a clear positive correlation only for Germany (see Appendix 1.6).

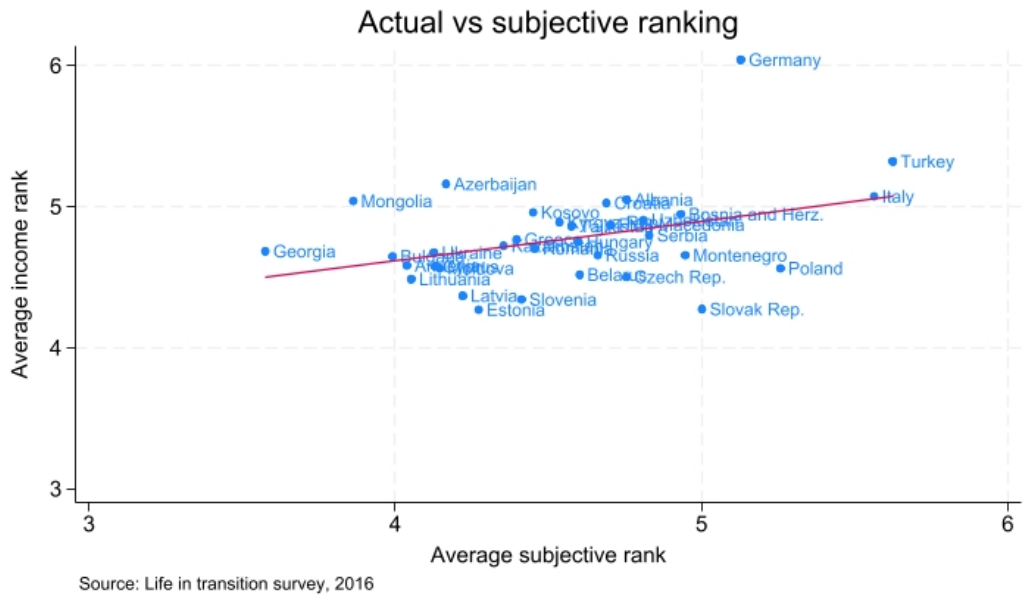


Figure 1.3: Correlation between actual and subjective income rank

Notes: The graph shows the correlation between the country's average subjective income rank (x-axis) and the country's average actual income rank (y-axis).

Figures 1.4 reports the distribution of the answers to the question on general trust by actual and subjective income decile. The distribution by actual income rank is pretty stable across deciles, while the distribution by subjective income rank varies much across deciles.

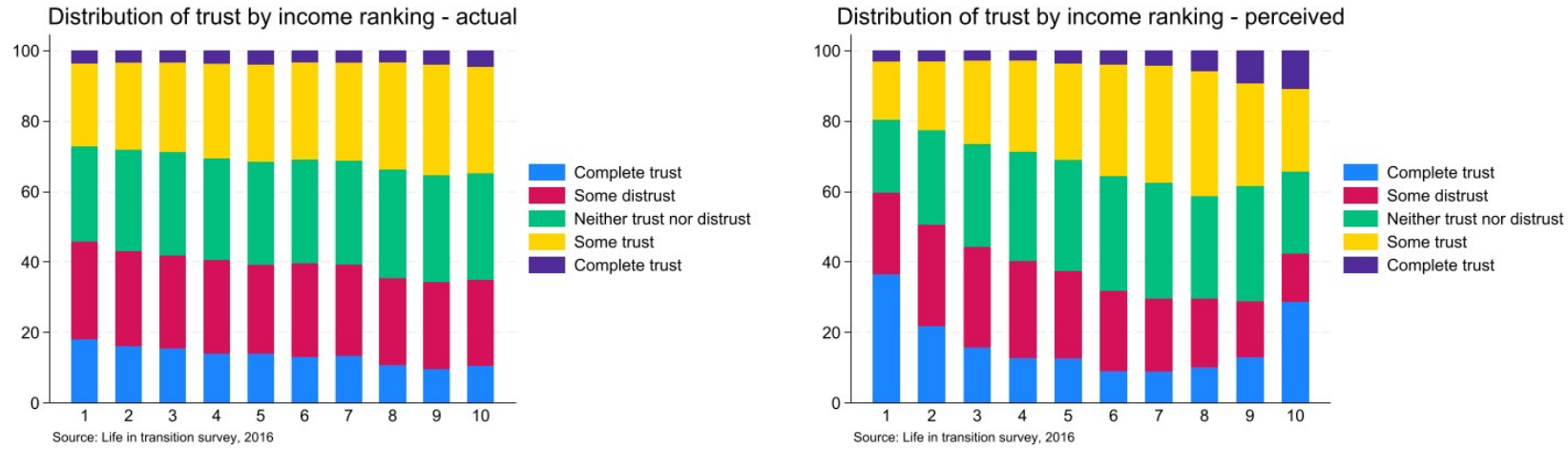


Figure 1.4: Distribution of trust by income rank

Notes: The graphs show the share of individual responses to the question *"Generally speaking, do you think that most people can be trusted or that you can't be too careful in dealing with people?"* by actual and subjective position on the income ladder.

Lastly, in Figure 1.5, the correlation between average trust and perception bias is reported at the country level. For each country, the share of people answering "some trust" or "complete trust" to the question on general trust[2] is calculated. At the same time, the perception bias is reported as the average difference between actual and subjective income rank in each country. As the figure shows, the correlation between trust and perception bias is almost null. However, Mongolia and Germany appear as outliers in the sample, reporting high levels of trust and high but negative levels of perception bias. Indeed, excluding these two countries makes the correlation clearly positive (Figure 1.6).

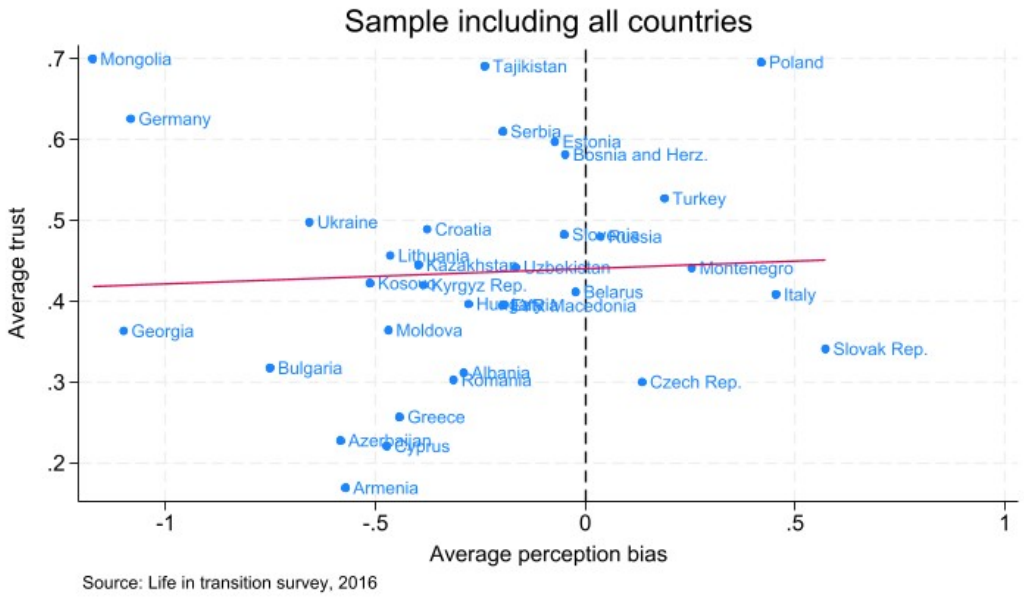


Figure 1.5: Correlation between average trust and perception bias (all countries)

Notes: The graphs show the correlation between countries' average perception bias in income rank (x-axis) and countries' share of people responding "some trust" or "complete trust" to the question *"Generally speaking, do you think that most people can be trusted or that you can't be too careful in dealing with people?"* (y-axis).

[2] *"Generally speaking, do you think that most people can be trusted or that you can't be too careful in dealing with people?"*

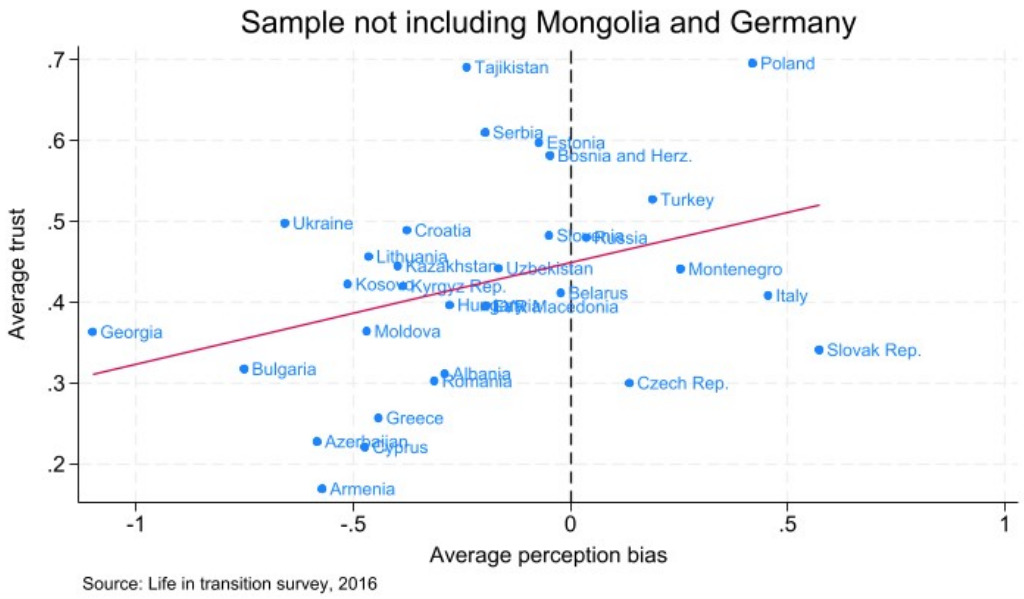


Figure 1.6: Correlation between average trust and perception bias (subsample)

Notes: The graphs show the correlation between countries' average perception bias in income rank (x-axis) and countries' share of people responding "some trust" or "complete trust" to the question *"Generally speaking, do you think that most people can be trusted or that you can't be too careful in dealing with people?"* (y-axis) for all the countries in the survey except for Mongolia and Germany.

The next section describes the empirical strategy used to study in depth relationship between trust measures and income ranks, showing that perceived economic success (and its bias) is a more sensitive indicator of variations in trust levels than actual economic success. After identifying the relationship between subjective income and trust, the perception bias is further investigated to understand the main determinants of its variation.

1.3 Empirical strategy

The empirical strategy follows three steps of specifications. First, subjective and actual income ranks are compared to understand which one helps better explaining variations in trust. Then, in the second step, the effect of the perception bias is isolated, to explore the relationship between the sign of the

bias and trust. Lastly, plausible determinants of the bias are explored. Results are reported in Section 3.4.

The first step of the analysis compares the two measures of economic success: actual vs subjective income rank. For this purpose, the following empirical models are under analysis:

$$trust_i = \alpha + \beta_1 \text{actual income rank}_i + \beta_2 X_{i,c,r} + \epsilon_i \tag{1.2}$$

$$trust_i = \alpha + \beta_1 \text{subjective income rank}_i + \beta_2 X_{i,c,r} + \epsilon_i \tag{1.3}$$

where i = individual, c = country and r = region.

As Alesina and La Ferrara (2002) show in their work on the determinants of trust, individual trust is influenced by traumatic experiences, discrimination, heterogeneity of the community and economic success in terms of income and education. For this reason, the following information is included in the analysis as relevant covariates $X_{i,c,r}$:

- proxies for recent history of traumatic experience: being a victim of prosecution, torture or acts of violence before 1989;
- proxies for belonging to a discriminated group: gender, ethnicity and religion;
- proxies for living in a highly heterogeneous community: regional fixed effects;
- additions: years in the city/village/town, education, age, height, weight.

In addition, another control is added for individual risk preferences that might affect trust or trustworthiness: the willingness to take risks. Indeed, according to Siegrist (2021), trust towards unknown people is related to the perceived riskiness of the person one needs to interact with. Hence, risk preferences can induce risk-loving people to be more willing to trust others than risk-averse people. The variable *willingness to take risks* collects the response

to the following question: *"Please, rate your willingness to take risks, in general, on a scale from 1 to 10, where 1 means that you are not willing to take risks at all, and 10 means that you are very much willing to take risks".*

The second step of the analysis focuses on the relationship between the type of bias and trust, hence separating the subjective income rank into two parts: actual income rank + perception bias. Here's the model under analysis:

$$trust_i = \alpha + \beta_1 \text{perception bias}_i + \beta_2 \text{actual income rank}_i + \beta_3 risk_i + \beta_4 X_{i,c,r} + \epsilon_i \tag{1.4}$$

where

- perception bias$_i$ $= \begin{cases} 1 & \text{negative} \\ 2 & \text{null} \\ 3 & \text{positive} \end{cases}$

Controls include regional fixed effects, age, education, gender, weight, height, being a victim of prosecution, torture or acts of violence before 1989, length of time in the city/village/town, religion, and ethnicity.

The last step of the empirical analysis aims at exploring the determinants of the perception bias. The following model is estimated:

$$\text{perception bias}_i = \alpha + \beta_1 \text{satisfaction}_i + \beta_2 \text{optimism}_i + \beta_3 \text{success}_i + \beta_4 \text{meet fam\&friends}_i + \beta_5 X_{i,r} + \epsilon_i \tag{1.5}$$

Controls include regional fixed effects, age, education, gender, weight, height, being a victim of prosecution, torture or acts of violence before 1989, religion, and ethnicity.

The variables highlighted in model (1.5) are used to proxy for overall satisfaction with the local economy, institutions and safety nets. In particular, two

variables help elicit people's perception of intergenerational mobility by asking about past economic upward scaling (satisfaction) and the future expectations (optimism). First, the level of satisfaction is proxied using the answer to the following question *"Do you agree to have done better in life than your parents?"*:

$$satisfaction_i = \begin{cases} 0 & \text{disagree or strongly disagree} \\ 1 & \text{agree or strongly agree} \end{cases}$$

Second, the level of optimism is proxied using the answer to the following question *"Do you agree that children who are born now will have a better life than my generation?"*:

$$optimism_i = \begin{cases} 0 & \text{disagree or strongly disagree} \\ 1 & \text{agree or strongly agree} \end{cases}$$

Instead, the variable *success* aims at measuring perceived institutional fairness. The variable reports the answer to the following question *"In your opinion, which of the following factors is the most important to succeed in life in your country now?"*:

$$success_i = \begin{cases} 1 & \text{effort and hard work} \\ 2 & \text{intelligence and skills} \\ 3 & \text{political connections} \\ 4 & \text{breaking the law} \end{cases}$$

The variables just described, as shown by Alesina, Stantcheva, and Teso (2018) and Benabou and Tirole (2006), affect individual beliefs and are corre-

lated to cognitive distortions about the state of the world.

Lastly, another variable is included, i.e. the network net available to interviewees. The variable *meet fam&friends* reports the answer to the following question *"How often do you meet up with friends or relatives who are not living with you in your household, on average?"*:

$$\text{meet fam\&friends}_i = \begin{cases} 1 & \text{on most days} \\ 2 & \text{once or twice a week} \\ 3 & \text{once or twice a month} \\ 4 & \text{less often than once a month} \\ 5 & \text{never} \end{cases}$$

In particular, the more frequent the encounters of family and friends outside the household, the more frequent the economic interactions, which can exacerbate or mitigate perception biases in economic status.

All these variables (*satisfaction, optimism, success and meet fam&friends*) are analysed to understand the mechanisms behind biased perceived economic status, which can be related to a lack of solid network ties, perceived unfairness in the institutions and perceived stagnation of the economy.

1.4 Results

The empirical specifications includes only the answers of the first respondent in the household, whom the interviewer randomly selected. Standard errors are clustered at the regional level in all regressions to control for within-region correlation.

As anticipated in Section 1.2, all trust measures are reported in a di-

chotomic format (1 = some or complete trust/0 = some or complete distrust), and neutral or undefined responses are excluded from the analysis. For this reason, the final number of observations included in the analysis is lower than the total number of households interviewed.

The coefficients for equations (1.2) and (1.3) are estimated using a linear probability model. Table 1.1 reports the coefficients results using, as the dependent variable for trust, the response to the question on general trust[3]. The coefficients from equations (1.2) and (1.3) are reported in columns (1) and (3). In columns (2) and (4), instead, the willingness to take risks is added as a control.

Results show that substituting the measure of actual rank with subjective rank increases the R^2 for trust by around 2 points. Moreover, while a one-step increase in the actual income rank explains a 3% variation in the measure of trust, a one-step increase in the perceived rank explains around 6% of the variation in the measure of trust. Adding the variable on the willingness to take risk reduces the coefficient magnitude of both income rank variables. However, the difference in magnitude between actual and subjective income rank is constant (around 3 percentage points), as well as the values of R^2.

As a robustness check, Table 1.2 shows the regression results for equation (1.3) using the other trust measures described in Section 1.2, i.e. trust towards family members, neighbours, foreigners and the expected likelihood of returned wallet. As expected, subjective income rank and willingness to take risks have a quasi-null correlation with trust in family members. At the same time, the coefficients are large and significant using trust towards newly met people, foreigners, and the likelihood of a returned wallet as dependent variables.

However, the R^2 reported in columns (1) and (2) are way below the values

[3] *"Generally speaking, do you think that most people can be trusted or that you can't be too careful in dealing with people?"*

shown in Table 1.1, suggesting that stronger network ties (family and neighbours) reduce the correlation between economic success and trust. Another explanation for the reduction in the explanatory power follows the intuition by Glaeser et al. (2000), who found that the response to the question of general trust is related to trustworthiness more than trust. Hence, differences in subjective rank might be more indicative of differences in individuals' trustworthiness (measured by the question on general trust) rather than trust in others (measured by the more targeted questions).

Table 1.3 reports the estimated coefficients of equation (1.4) using a linear probability model. Results show that a positive bias correlates with a higher level of trust. Therefore, people with a higher perceived economic status are also more likely to trust others. Interestingly, the coefficient for the negative bias changes sign when adding the variables *actual income rank* to the equation and the magnitude of the correlation between the perception bias and trust increases. On the other hand, the willingness to take risks does not capture significant variations.

Coefficients of equation (1.5) are estimated using an ordered probit model, and results are reported in Table 1.4. The pseudo R^2 of all the probit regressions is around 0.4, indicating a high predicting power of the regressors included. Moreover, as Table 1.4 shows, adding country-fixed effects does not significantly affect the regression results. To facilitate the interpretation of the coefficients, margins are reported in Table 1.5.

As Table 1.5 shows, the variables *satisfaction*, *success* and *meet fam&friends* are more significantly correlated to the perception bias, while the variable *optimism* does not capture significant variations. The coefficients indicate a higher likelihood for individuals with optimistic perspectives for their children (*optimism*) to have a positive bias. On the other hand, the more sparse the

reported interactions with family and friends outside the household, the more likely to have a negative bias. Similarly, the more people think success is not linked to skills or effort, the higher the likelihood of having a negative bias[4]. Interestingly, people that believe to have been more successful in life compared to their parents (*success*), are also more likely to report a negative bias.

To conclude, perception about intergenerational mobility and fairness of institutions, and tightness of network links strongly correlate with variations in perception biases about one's economic success. In particular, loose network links and lack of institutional trust are more likely to associate with a negative bias in perceived economic success. At the same time, a negative bias in perceived economic success is linked to a lower trust in a society.

[4]The coefficients for the variable success are reported in relationship to the base value $success$ = effort and hard work.

VARIABLES	(1) General trust	(2) General trust	(3) General trust	(4) General trust
actual income rank	0.025***		0.021***	
	(0.004)		(0.004)	
subjective income rank		0.059***		0.053***
		(0.005)		(0.005)
willingness to take risks			0.02***	0.018***
			(0.002)	(0.002)
Constant	0.37***	0.354***	0.241***	0.232***
	(0.069)	(0.069)	(0.072)	(0.071)
Observations	24,139	29,342	23,609	28,675
R-squared	0.164	0.180	0.174	0.188
Controls	yes	yes	yes	yes
FE	region	region	region	region

Table 1.1: Comparison between actual and subjective income rank

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Columns (1) and (3) report the estimated coefficients for equation (1.2) using a linear probability model. In column (3), the control *"willingness to take risk"* is added. Columns (2) and (4) report the estimated coefficients for equation 1.3 using linear regressions. In column (4), the control *"willingness to take risk"* is added. *Actual income rank* is measured as the decile of belonging of each respondent to the survey. *Subjective income rank* is measured as the survey response to the question *"Imagine a ten-step ladder where on the bottom, the first step, stand the poorest 10% people in our country, and on the highest step, the tenth, stand the richest 10% people in your country. On which step of the ten is your household today?"*. Both variables for income rank are standardised. *Willingness to take risk* is measured as the survey response to the question *Please, rate your willingness to take risks, in general, on a scale from 1 to 10, where 1 means that you are not willing to take risks at all, and 10 means that you are very much willing to take risks.*. Controls include being a victim of violence before 1989, gender, ethnicity, religion, years living in the city, education, age, height and weight. Regional-fixed effects are included in each specification. Country-fixed effects are included in each specification. Standard errors are clustered at the regional level. Coefficients on the other covariates are reported in the Appendix 1.6.

VARIABLES	(1) Trust family	(2) Trust neighbours	(3) Trust newly met ppl	(4) Trust foreigners	(5) Wallet returned
subjective income rank	0.003***	0.021***	0.028***	0.018***	0.031***
	(0.001)	(0.003)	(0.005)	(0.004)	(0.004)
willingness to take risks	0.0002	-0.002	0.012***	0.011***	0.006***
	(0.0004)	(0.001)	(0.002)	(0.002)	(0.002)
Constant	1.001***	0.746***	0.094	0.08	0.26***
	(0.036)	(0.061)	(0.083)	(0.076)	(0.066)
Observations	37,862	34,472	27,968	27,781	40,716
R-squared	0.060	0.078	0.146	0.136	0.126
Controls	yes	yes	yes	yes	yes
FE	region	region	region	region	region

Table 1.2: Correlation between subjective income rank and trust measures

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Table 2 reports the estimated coefficients for equation (1.3) using a linear probability model. In columns (1)-(4), the dependent variables collect the responses to the following questions: column (1) for *"To what extent do you trust the family living with you?"*; column (2) for *"To what extent do you trust people from your neighbourhood?"*; column (3) for *"To what extent do you trust people you meet for the first time?"*; column (4) for *"To what extent do you trust foreigners?"*. The dependent variables are equal to 1 if the response to the question is *"some trust"* or *"complete trust"* and equal to 0 otherwise. Instead, in column (5), the dependent variable is equal to 1 if the response to the question *"Suppose you lost your (purse/wallet) containing your address details, and it was found in the street by someone living in this neighbourhood. How likely would it be returned to you with nothing missing?"* is *"very likely"* or *"quite likely"*. *Subjective income rank* is measured as the survey response to the question *"Imagine a ten-step ladder where on the bottom, the first step, stand the poorest 10% people in our country, and on the highest step, the tenth, stand the richest 10% people in your country. On which step of the ten is your household today?"*. The variable for income rank is standardised. *Willingness to take risk* is measured as the survey response to the question *Please, rate your willingness to take risks, in general, on a scale from 1 to 10, where 1 means that you are not willing to take risks at all, and 10 means that you are very much willing to take risks.* Controls include being a victim of violence before 1989, gender, ethnicity, religion, years living in the city, education, age, height and weight. Regional-fixed effects are included in each specification. Standard errors are clustered at the regional level. Coefficients on the other covariates are reported in the Appendix 1.6.

VARIABLES	(1) General trust	(2) General trust	(3) General trust
bias = negative	0.020**	-0.046***	-0.044***
	(0.009)	(0.012)	(0.012)
bias = positive	0.023**	0.063***	0.058***
	(0.011)	(0.012)	(0.012)
actual income rank		0.064***	0.058***
		(0.007)	(0.007)
willingness to take risks			0.019***
			(0.002)
Constant	0.322***	0.372***	0.248***
	(0.068)	(0.070)	(0.072)
Observations	29,740	24,139	23,609
R-squared	0.167	0.168	0.178
Controls	yes	yes	yes
FE	region	region	region

Table 1.3: Correlation between trust and perception bias

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Table 3 reports the coefficients from equation (1.4) estimated using a linear probability model. The dependent variable is a binary variable equal to 1 if the response to the question *"Generally speaking, do you think that most people can be trusted or that you can't be too careful in dealing with people?"* was *complete trust* or *some trust*, equal to 0 otherwise. The coefficients for the perception bias (negative and positive) are reported relative to the baseline level of null perception bias. *Actual income rank* is measured as the decile of belonging of each respondent to the survey. The measure for income rank is standardised. *Willingness to take risk* is measured as the survey response to the question *Please, rate your willingness to take risks, in general, on a scale from 1 to 10, where 1 means that you are not willing to take risks at all, and 10 means that you are very much willing to take risks.*. Controls include being a victim of violence before 1989, gender, ethnicity, religion, years living in the city, education, age, height, and weight. Regional-fixed effects are included in each specification. Regional-fixed effects are included in each specification. Standard errors are clustered at the regional level. Coefficients on the other covariates are reported in the Appendix 1.6.

VARIABLES	(1) Perception bias	(2) Perception bias
satisfaction	-0.087***	-0.068***
	(0.025)	(0.026)
optimism	0.023	0.041*
	(0.023)	(0.024)
success = intelligence and skills	0.002	-0.003
	(0.024)	(0.024)
success = political connections	-0.055*	-0.067**
	(0.029)	(0.028)
success = breaking the law	-0.038	-0.043
	(0.039)	(0.040)
meet fam&friends	-0.028**	-0.024**
	(0.012)	(0.011)
Observations	24,234	24,234
Controls	yes	yes
FE	no	region

Table 1.4: The determinants of perception bias

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Using an ordered probit model, Table 4 reports the estimated coefficient from equation (1.5). The dependent variable *Perception bias* equals 1 if the bias is negative, 2 if null and 3 if positive. The variable *satisfaction* equals 1 if the response to the question *Do you agree to have done better in life than your parents?* is *agree* or *strongly agree*, equals 0 otherwise. The variable *optimism* equals 1 if the response to the question *Do you agree that children who are born now will have a better life than my generation?* is *agree* or *strongly agree*, equals 0 otherwise. The variable *success* collects the responses to the question *In your opinion, which of the following factors is the most important to succeed in life in your country now?*, the responses are ranked increasingly in negative factors and reported in relationship to the value *success* = effort and hard work. The variable *meet fam&friends* collects the responses to the question *How often do you meet up with friends or relatives who are not living with you in your household?*, and the responses are ranked decreasingly in frequency. Controls include age, gender, height, weight, education, being a victim of prosecution before 1989, religion and ethnicity. Column (1) reports the results without fixed effects, and column (2) reports the results with regional fixed effects. Standard errors are clustered at the regional level. Coefficients on the other covariates are reported in the Appendix 1.6.

VARIABLES	(1) satisfaction	(2) optimism	(3) success	(4) meet fam&friends
negative bias	0.023***	-0.015*	0.008**	0.008**
	(0.009)	(0.008)	(0.004)	(0.004)
null bias	0.001***	-0.001*	0.0004**	0.0004**
	(0.001)	(0.0004)	(0.0002)	(0.0002)
positive bias	-0.024***	0.015*	-0.008**	-0.009**
	(0.009)	(0.009)	(0.004)	(0.004)
Observations	24,234	24,234	24,234	24,234

Table 1.5: Marginal effects of coefficients in Table 1.4

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Table 5 reports the margins of the estimated coefficient from equation (1.5) using an ordered probit model. The variable *satisfaction* equals 1 if the response to the question *Do you agree to have done better in life than your parents?* is *agree* or *strongly agree*, equals 0 otherwise. The variable *optimism* equals 1 if the response to the question *Do you agree that children who are born now will have a better life than my generation?* is *agree* or *strongly agree*, equals 0 otherwise. The variable *success* collects the responses to the question *In your opinion, which of the following factors is the most important to succeed in life in your country now?*, the responses are ranked increasingly in negative factors. The variable *meet fam&friends* collects the responses to the question *How often do you meet up with friends or relatives who are not living with you in your household?*, and the responses are ranked decreasingly in frequency. Controls include age, gender, height, weight, education, being a victim of prosecution before 1989, religion and ethnicity. Regional-fixed effects are included. Standard errors are clustered at the regional level.

1.5 Conclusions

This first chapter explores the effects of biases in perceived economic success on individuals' trust. The analysis is based on survey data from the LiTS administered in 2016 by the EBRD and the World Bank to over 50,000 households in 34 countries.

The study begins by describing the survey responses on trust and income status, highlighting how each variable interacts with the others. Then, regression models are run to test the levels of correlation between perception biases in income status and trust. Lastly, an exploration of the possible determinants of the bias is implemented by studying the correlation with other salient survey answers.

Results find that subjective economic success explains variations in trust better than actual economic success. In particular, negative biases in subjective economic status are linked to lower trust in others. Moreover, satisfaction about intergenerational mobility, perceived corruption in institutions, and tightness of network ties are highly correlated with biases in perceived economic status. In particular, loose network ties and low trust in institutions are associated with a more negative perception bias and, consequently, lower trust among people.

The analysis carries exciting insights into the relationship between trust and economic success and it has a high external validity, as the study includes 34 countries across Europe and Asia. However, it also has some weaknesses worth improving in future studies. The main one is that the analysis is based on survey data, a method that could influence the responses to the questions on trust. Indeed, eliciting trust through laboratory experiments, as shown by Berg et al. (1995) or Glaeser et al. (2000), can give more precise estimates than survey questions. Moreover, the lack of exogenous variation in the sample

selection leaves no space for causal inference. It is left to future explorations to identify the causal link between perception biases in economic status and trust.

1.6 Appendix

Regressions controls coefficients

VARIABLES	(1) General trust	(2) General trust	(3) General trust	(4) General trust
age	0.0004*	0.0003	0.001***	0.001***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
education	0.023***	0.018***	0.021***	0.016***
	(0.003)	(0.003)	(0.003)	(0.003)
gender	-0.003	-0.005	0.007	0.004
	(0.007)	(0.006)	(0.007)	(0.006)
weight	0.0002*	0.0002	0.0001	10e-05
	(0.0001)	(10e-05)	(0.0001)	(10e-05)
height	-6e-05	-2e-05	-4e-05	-2e-05
	(7e-05)	(7e-05)	(7e-05)	(7e-05)
victim	0.064*	0.052*	0.050	0.038
	(0.034)	(0.029)	(0.034)	(0.030)
years in town	9e-05	0.0001	9e-05	10e-05
	(10e-05)	(9e-05)	(10e-05)	(9e-05)
Observations	24,139	29,342	23,609	28,675
R-squared	0.164	0.180	0.174	0.188
Other controls	religion, ethnicity	religion, ethnicity	religion, ethnicity	religion, ethnicity
FE	region	region	region	region

Table 1.6: Comparison between actual and subjective income rank: controls coefficients

VARIABLES	(1) Trust family	(2) Trust neighbours	(3) Trust newly met ppl	(4) Trust foreigners	(5) Wallet returned
age	-6e-05	0.001***	0.002***	0.001***	0.001***
	(5e-05)	(0.0002)	(0.0003)	(0.0002)	(0.0002)
education	0.002**	0.008***	0.013***	0.017***	0.002
	(0.001)	(0.002)	(0.003)	(0.002)	(0.002)
gender	0.001	-0.008*	-0.004	-0.013**	-0.007
	(0.001)	(0.004)	(0.006)	(0.006)	(0.006)
weight (kg)	4e-05	0.0001*	4e-05	-4e-05	2e-05
	(4e-05)	(7e-05)	(9e-05)	(8e-05)	(9e-05)
height (cm)	-4e-05	4e-06	-4e-05	-1e-05	0.0001
	(3e-05)	(5e-05)	(6e-05)	(5e-05)	(7e-05)
victim	-0.044***	0.008	0.095**	0.133***	0.136***
	(0.015)	(0.023)	(0.046)	(0.041)	(0.043)
years in town	-2e-05	-8e-06	10e-05	3e-05	-0.0001*
	(2e-05)	(5e-05)	(8e-05)	(7e-05)	(7e-05)
Observations	37,862	34,472	27,968	27,781	40,716
R-squared	0.060	0.078	0.146	0.136	0.126
Other controls	religion, ethnicity	religion, ethnicity	religion, ethnicity	religion, ethnicity	ethnicity
FE	region	region	region	region	region

Table 1.7: Correlation between subjective income rank and trust measures: controls coefficients

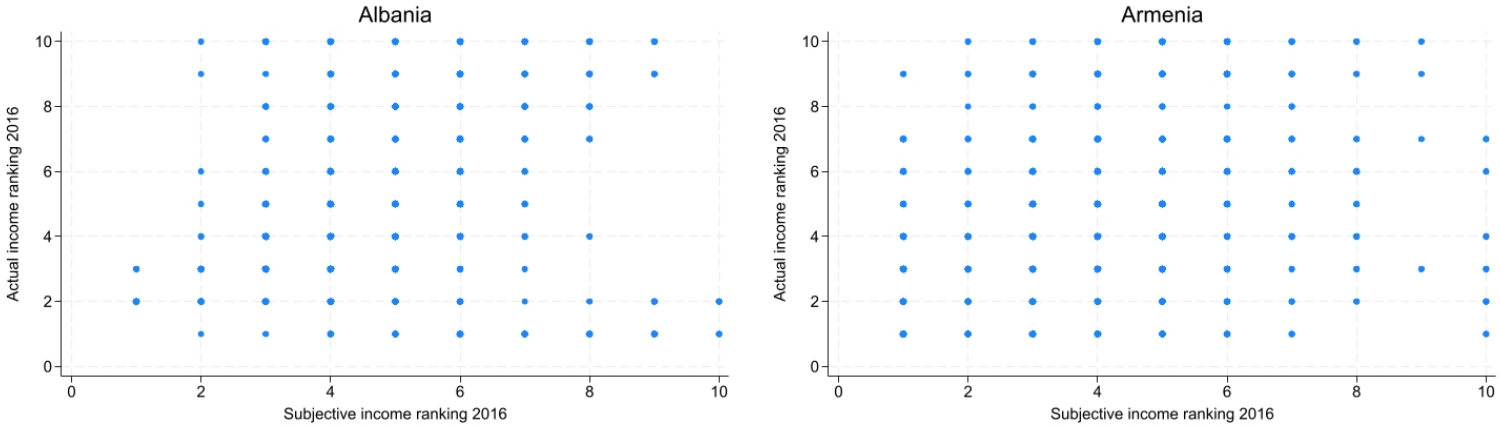
VARIABLES	(1) General trust	(2) General trust	(3) General trust
age	-7.13e-05	0.0004*	0.001***
	(0.0002)	(0.0002)	(0.0002)
education	0.025***	0.021***	0.020***
	(0.003)	(0.003)	(0.003)
gender	-0.007	-0.004	0.006
	(0.006)	(0.006)	(0.007)
weight (kg)	0.0002*	0.0002*	0.0001
	(9.68e-05)	(0.0001)	(0.0001)
height (cm)	-2.12e-05	-5.07e-05	-3.40e-05
	(6.51e-05)	(7.23e-05)	(7.24e-05)
victim	0.052*	0.065*	0.051
	(0.030)	(0.033)	(0.033)
years in town	0.0001	9.25e-05	8.64e-05
	(9.49e-05)	(9.69e-05)	(9.50e-05)
Observations	29,740	24,139	23,609
R-squared	0.167	0.168	0.178
Other controls	religion, ethnicity	religion, ethnicity	religion, ethnicity
FE	region	region	region

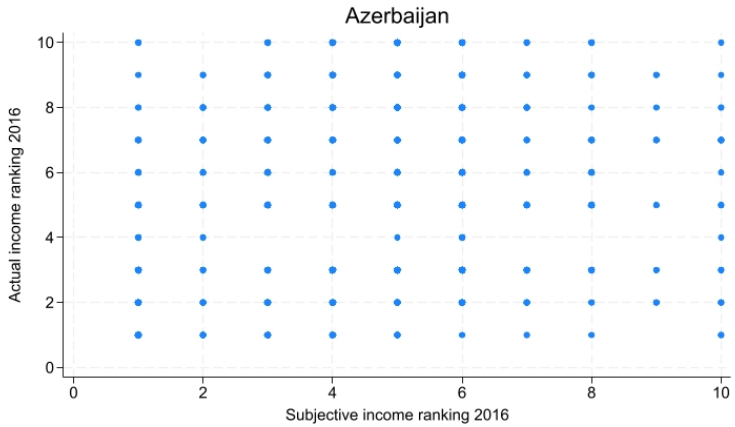
Table 1.8: Correlation between trust and perception bias: controls coefficients

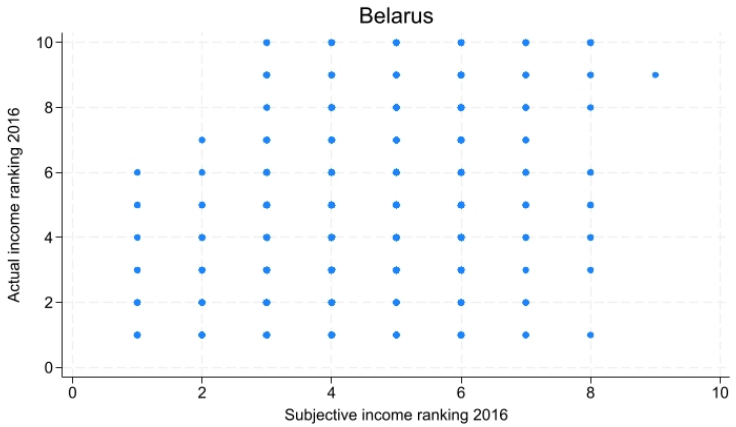
VARIABLES	(1) Perception bias	(2) Perception bias
age	0.006***	0.007***
	(0.001)	(0.001)
education	-0.102***	-0.094***
	(0.008)	(0.008)
gender	0.103***	0.123***
	(0.020)	(0.020)
weight (kg)	-0.001***	-0.001***
	(0.0003)	(0.0003)
height (cm)	-0.001***	-0.001***
	(0.0002)	(0.0002)
victim	0.136*	0.127
	(0.083)	(0.088)
Observations	24,234	24,234
Other controls	religion, ethnicity	religion, ethnicity
FE	region	region

Table 1.9: The determinants of perception bias: controls coefficients

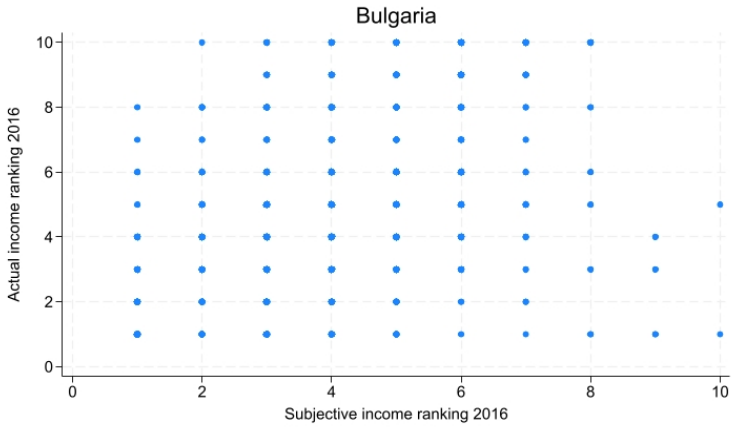
Correlation between subjective income ranking and acual income ranking by country

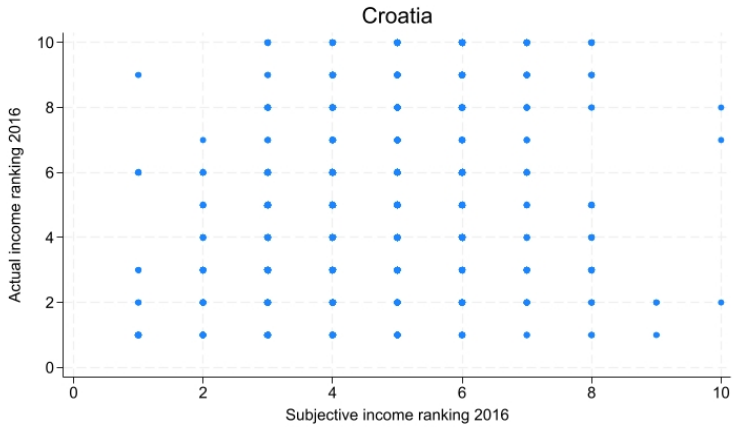


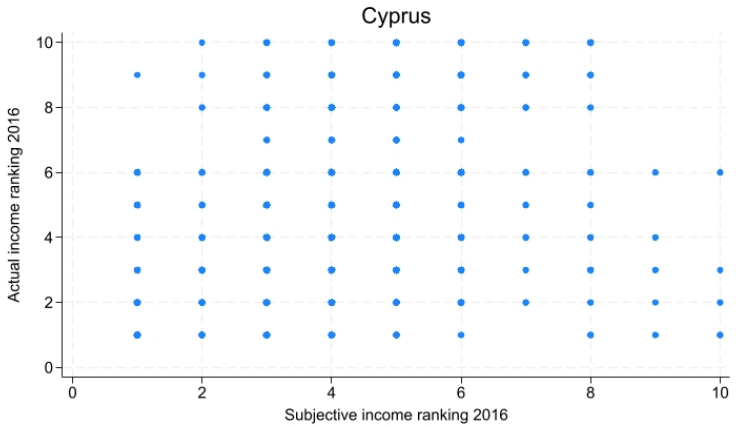

Azerbaijan
Actual income ranking 2016
Subjective income ranking 2016
0
2
4
6
8
10


Belarus
Actual income ranking 2016
Subjective income ranking 2016
0
2
4
6
8
10

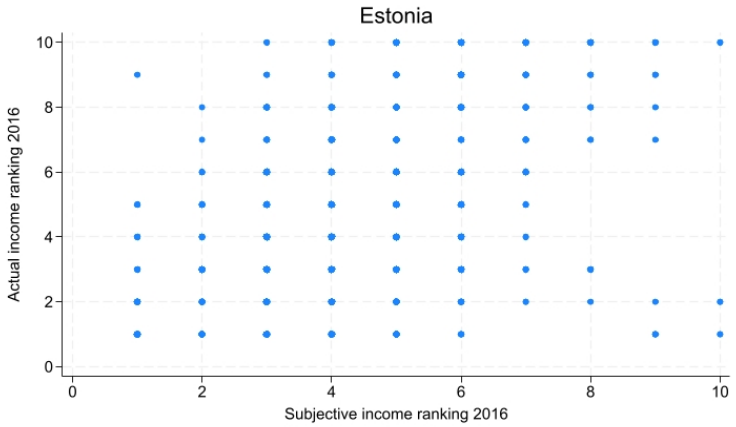

Bosnia and Herz.
Actual income ranking 2016
Subjective income ranking 2016
0
2
4
6
8
10

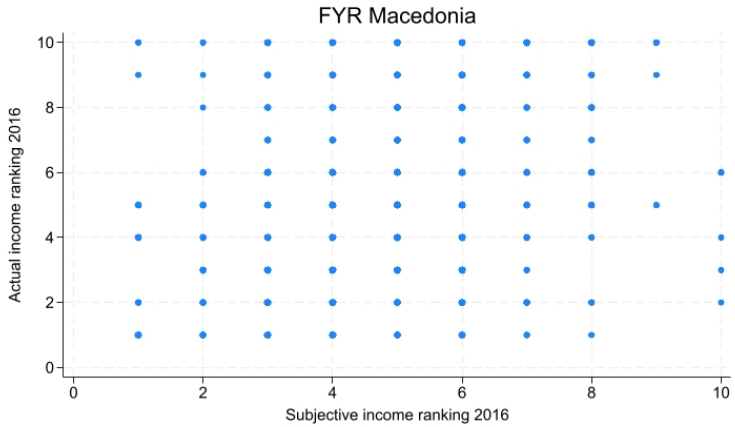

Bulgaria
Actual income ranking 2016
Subjective income ranking 2016
0
2
4
6
8
10

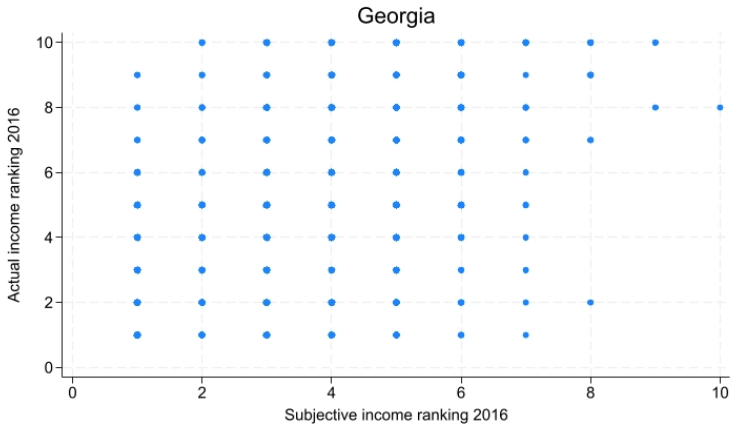

Croatia
Actual income ranking 2016
Subjective income ranking 2016
0
2
4
6
8
10

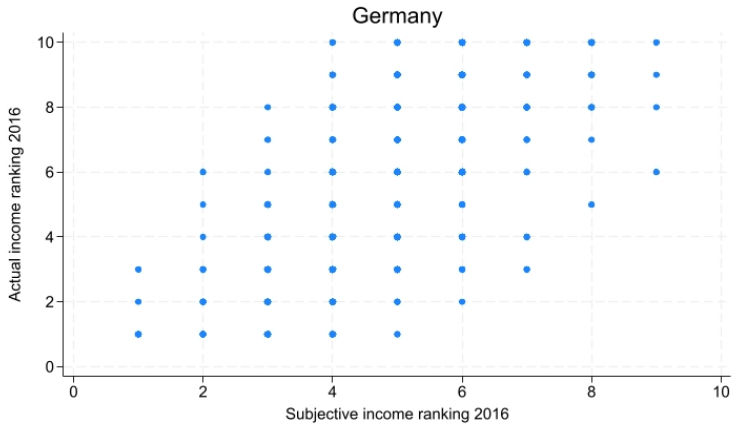

Cyprus
Actual income ranking 2016
Subjective income ranking 2016
0
2
4
6
8
10

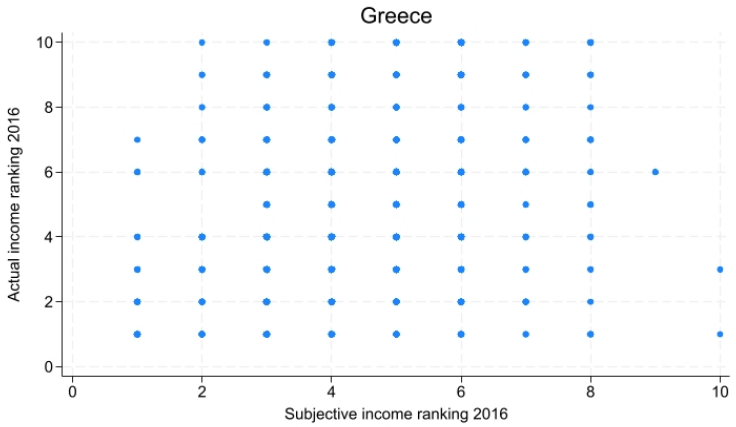

Czech Rep.
Actual income ranking 2016
Subjective income ranking 2016
0
2
4
6
8
10

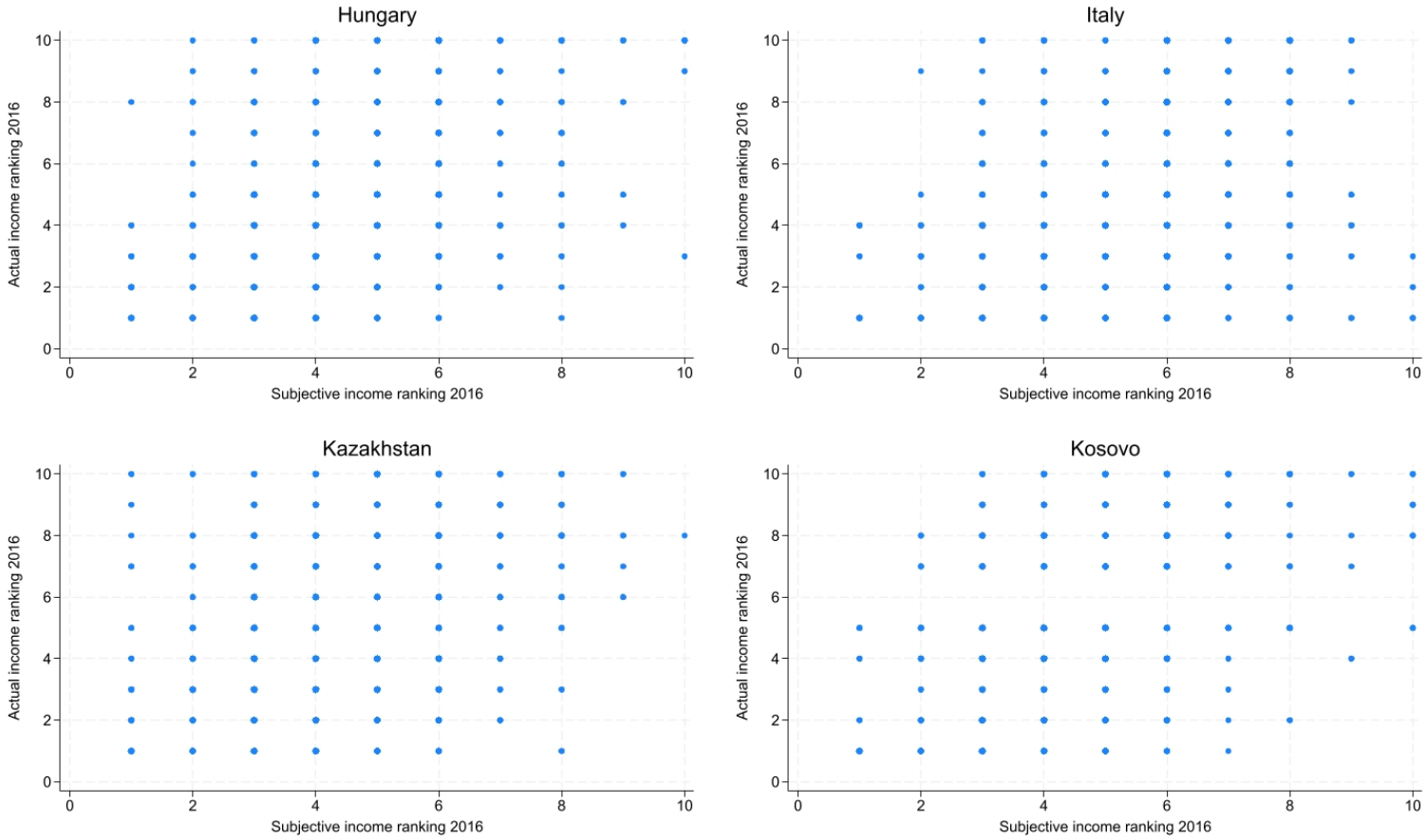

Estonia
Actual income ranking 2016
Subjective income ranking 2016
0
2
4
6
8
10

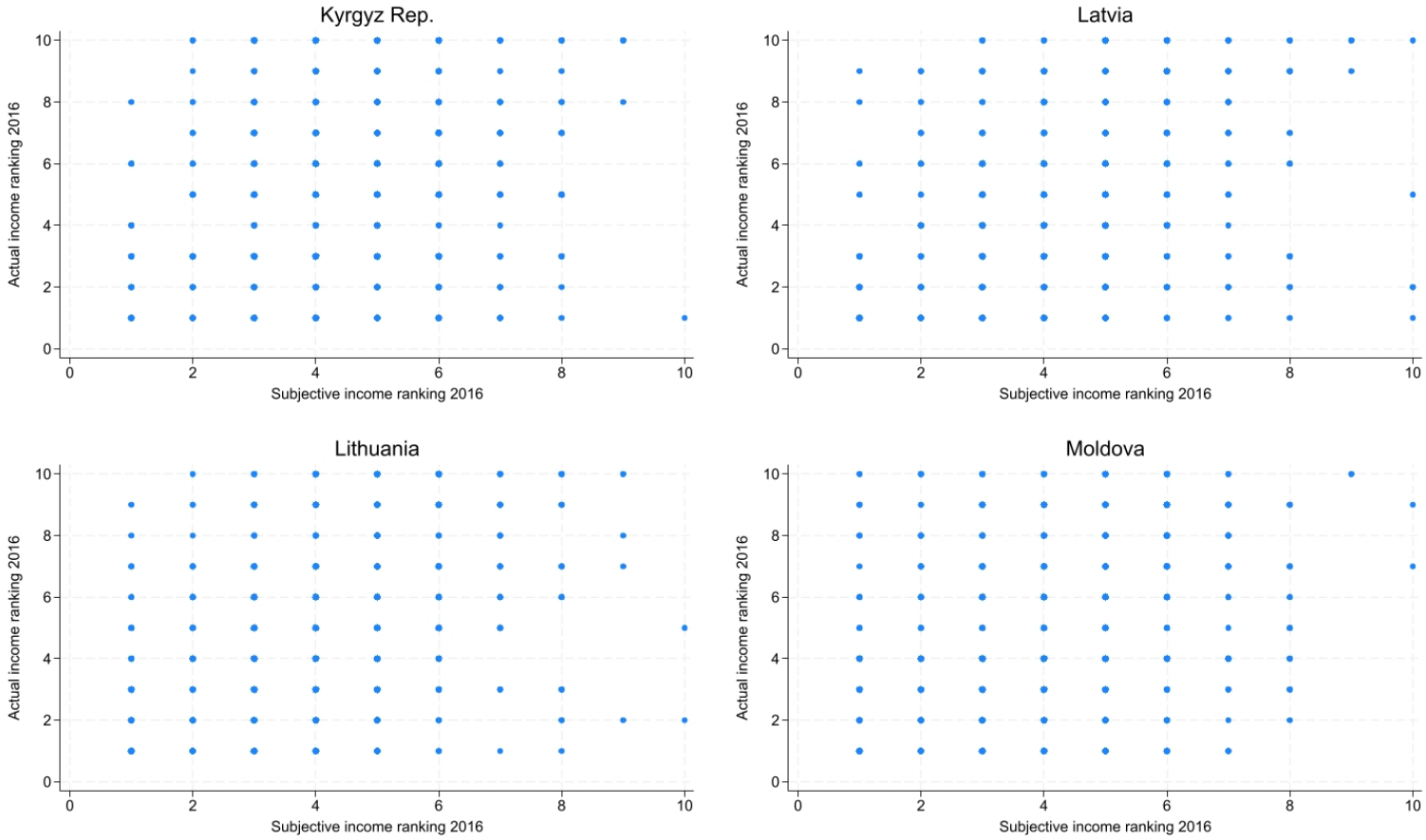

FYR Macedonia
Actual income ranking 2016
Subjective income ranking 2016
0
2
4
6
8
10

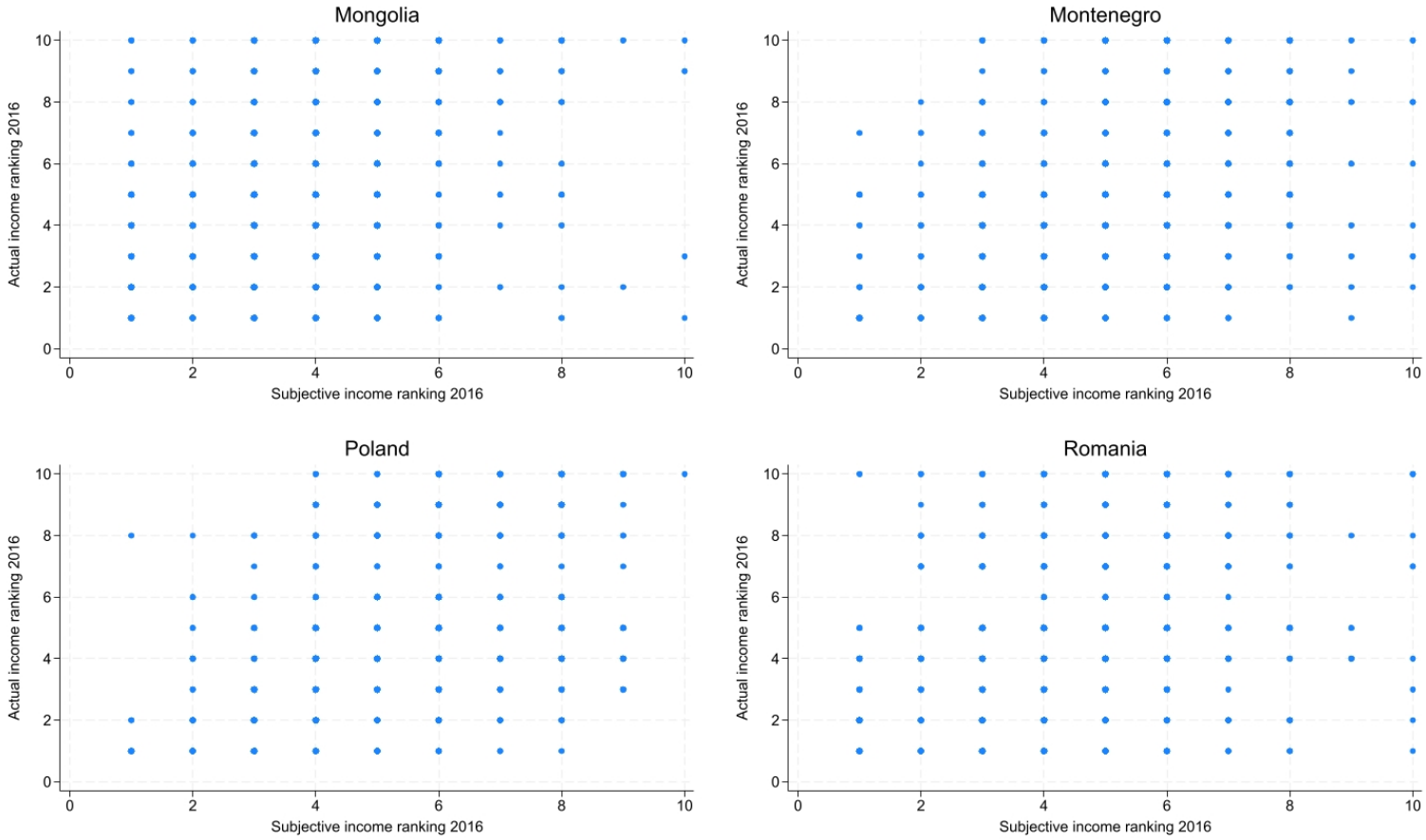

Georgia
Actual income ranking 2016
Subjective income ranking 2016
0
2
4
6
8
10

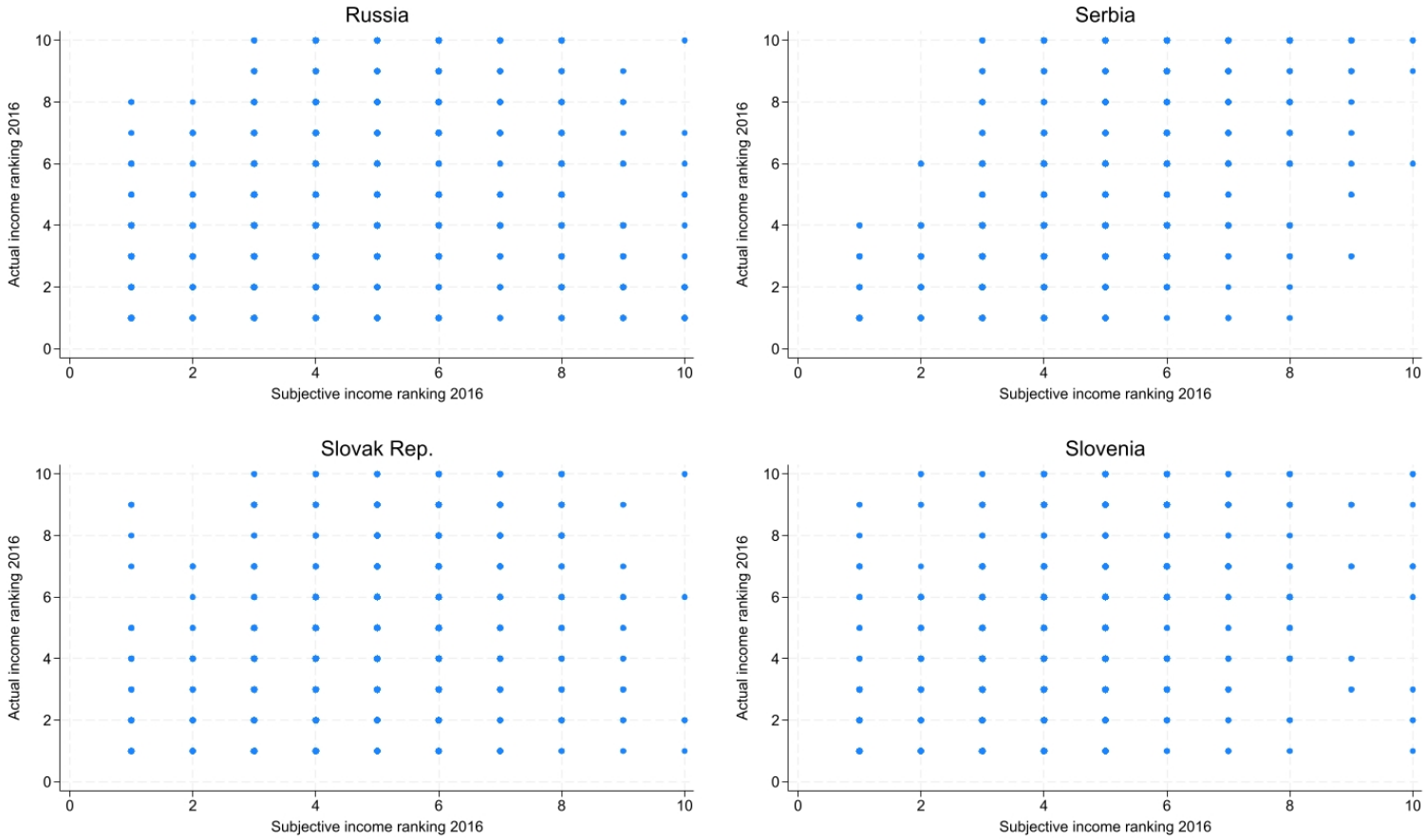

Germany
Actual income ranking 2016
Subjective income ranking 2016
0
2
4
6
8
10

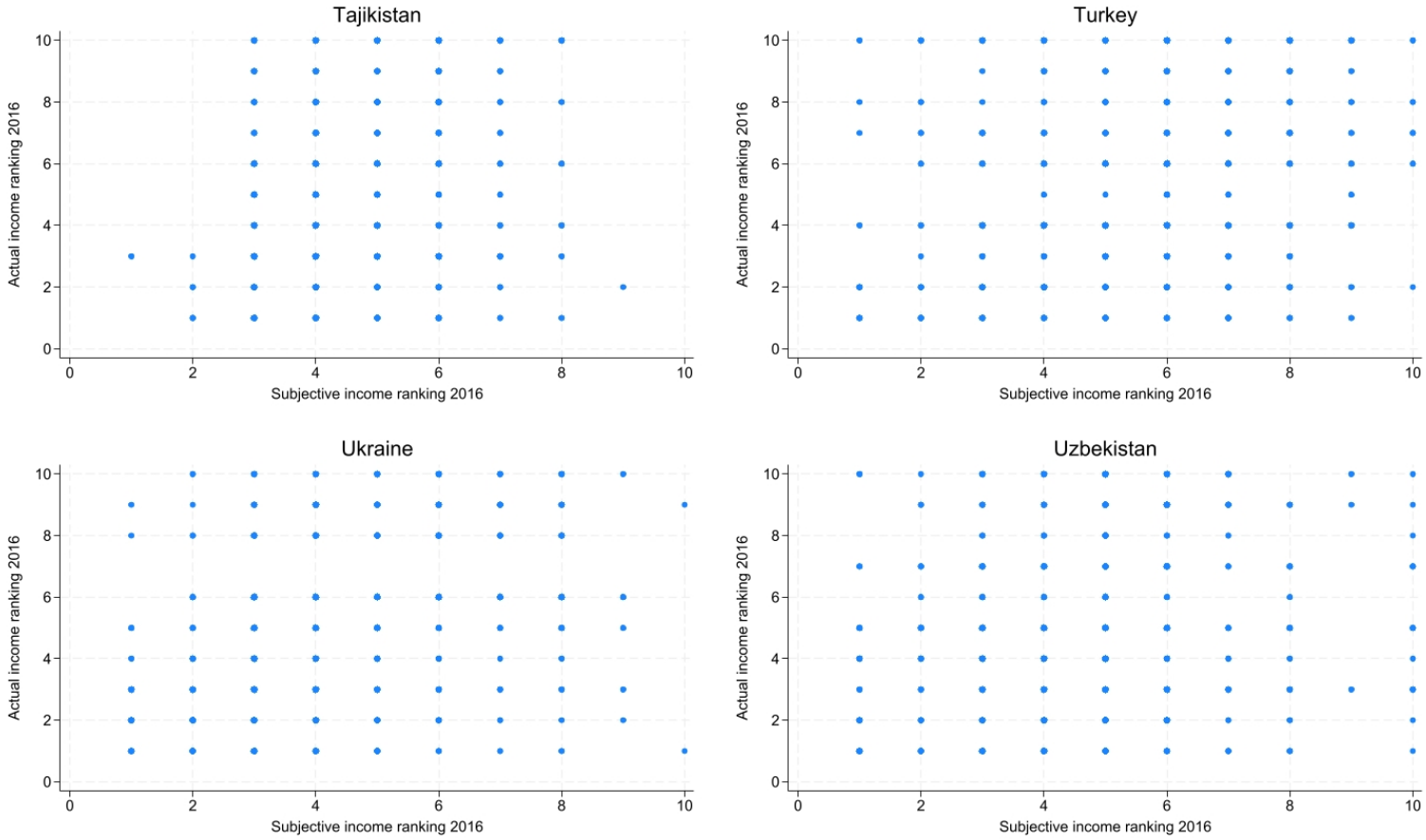

Greece
Actual income ranking 2016
Subjective income ranking 2016
0
2
4
6
8
10


Hungary
Italy
Kazakhstan
Kosovo
Actual income ranking 2016
Subjective income ranking 2016
0
2
4
6
8
10


Kyrgyz Rep.
Latvia
Lithuania
Moldova
Actual income ranking 2016
Subjective income ranking 2016
0
2
4
6
8
10


Mongolia
Montenegro
Poland
Romania
Actual income ranking 2016
Subjective income ranking 2016
0
2
4
6
8
10


Russia
Serbia
Slovak Rep.
Slovenia
Actual income ranking 2016
Subjective income ranking 2016
0
2
4
6
8
10


Tajikistan
Turkey
Ukraine
Uzbekistan
Actual income ranking 2016
Subjective income ranking 2016
0
2
4
6
8
10

Chapter 2

Risk-sharing motives

2.1 Introduction

The economic literature has broadly investigated the issue of enforcing contracts in informal settings, theoretically and empirically. This is not only relevant in developing countries, where income shocks are frequent, and insurance is primarily informal and incomplete (Robinson (2012)), but it also has a broader relevance, as the study of informal contracts enforcement helps to understand the underlying forces of cooperation, even in terms of culture and social norms (e.g. Berg et al. (1995), Fafchamps (2006)).

Among the possible realisations of informal contracts, intranetwork transfers (e.g. Foster and Rosenzweig (2001), Altonji, Hayashi, and Kotlikoff (1992)) and risk-sharing (e.g. Townsend (1995), Robinson (2012)) are the most commonly studied in the context of emerging markets. In particular, the economic literature considers risk-sharing a fundamental risk coping mechanism for the rural poor (Fafchamps and Gubert (2007), Townsend (1995)).

As Fafchamps and Gubert (2007) show, benefits from sharing risk are largest when individuals have uncorrelated incomes. However, the literature

agrees that empirically informal risk-sharing is often incomplete or asymmetric and it is mainly observed within close networks of households and friends (e.g. Fafchamps and Gubert (2007), Bramoulle and Kranton (2007), Townsend (1995)).

This work aims to contribute to the literature on risk-sharing by studying the motivations behind it. To my knowledge, no economic studies have yet investigated which are the most relevant channels underlying the decision to share risk. Nevertheless, citing Batson (2010): *"if one is interested in knowing more generally when and where help can be expected, and how effective it is likely to be [...] then understanding the underlying motivation is crucial"*.

The literature highlights two plausible mechanisms underlying risk-sharing and cooperation in emerging markets. On the one hand, following Arrow (1974), many argue that trust is fundamental in every economic interaction and more so in risk-sharing arrangements (e.g. Attanasio, Barr, Cardenas, Genicot, and Meghir (2012), Siegrist (2021), Evans and Krueger (2009)). On the other hand, some have shown how people engage in risk-sharing activities more often within networks creating social costs from non-cooperation (e.g. Altonji et al. (1992), Jackson et al. (2012), Jakiela and Ozier (2015)).

This work relies on a simple intuition: by observing people's behaviour in standard cooperative settings (in this case, public goods games), it is possible to identify individuals with collective preferences. The more often this population subset engages in risk-sharing activities, the more trust and trustworthiness are relevant in explaining risk-sharing group formation. On the other hand, social costs are more likely to motivate the phenomenon if collective preferences appear less relevant than network ties in predicting risk-sharing arrangements. The specific hypotheses tested are the following.

The first hypothesis is that risk-sharing is motivated by social

costs. Since these costs are highest within networks (Altonji et al. (1992)), one should observe a lower likelihood of forming risk-sharing groups outside networks.

The second hypothesis is that the motivation underlying risk-sharing is reciprocal trust. In this case, risk-sharing can also be achieved outside networks in communities that usually rely on cooperation in different settings. Indeed, when dealing with a person outside one's network, the likelihood that they are trustworthy is high in cooperative municipalities.

Intuitively, the two hypotheses can complement each other and hold simultaneously. Indeed, even within networks, one may observe group formation based on individual characteristics of trustworthiness. The objective of this study is, hence, to explore, theoretically and empirically, the two mechanisms and find which one prevails over the other, if any.

In this work, communities showing high levels of overall cooperation are identified as socially cohesive, following Chan, To, and Chan (2006) definition of social cohesion as: *"a state of affairs concerning both the vertical and the horizontal interactions among members of society as characterised by a set of attitudes and norms that includes trust, a sense of belonging and the willingness to participate and help, as well as their behavioural manifestations"*.

The study relies on a lab-in-the-field experiment implemented in Colombia in 2006, with over 2,000 participants from 70 municipalities observed interacting in cooperative games (Attanasio et al. (2012)). Participants in the experiment are asked first to play public goods games with randomly selected groups and, subsequently, risk-sharing games with endogenously chosen groups.

In the laboratory setting, risk-sharing groups are formed both within and outside existing networks, and full contribution is reached in around 90% of the groups. For this reason, collected data allow for an in-depth analysis of the

individual and social characteristics that might influence group formation in risk-sharing. Theoretical predictions from Nash Equilibria are used to support the empirical results.

Results show that, among individuals engaging in a risk-sharing group, coming from the same network increases the probability of being in the same risk-sharing group. However, within networks, risk-sharing groups are formed assortatively based on individual characteristics of trustworthiness. Moreover, outside networks, the likelihood of joining a risk-sharing group is strongly correlated with the level of social cohesion in the municipality, and groups formed in more cohesive municipalities are also larger on average. Interestingly, risk attitudes seem irrelevant in explaining group formation, both within and outside networks.

To conclude, high levels of trust in a society can increase the individual willingness to engage in risk-sharing activities and, following the predictions by Fafchamps and Gubert (2007), it can also increase the efficiency of risk-sharing by connecting individuals whose incomes are likely uncorrelated (outside the networks of family and friends).

The rest of the chapter is structured as follows: first, the laboratory experiment is described together with relevant summary statistics; then, theoretical predictions are analysed to study the incentives and disincentives for individuals to join risk-sharing groups; lastly, the empirical specification and its results are reported, comparing theoretical findings with the experimental data.

2.2 Laboratory experiment

The empirical data analysed in this chapter are collected through a laboratory experiment with around 2,300 participants from 70 Colombian municipalities.

See Attanasio et al. (2012) for information on the design of the experiment. Moreover, Chapter 3 of this dissertation includes furhter details on the characteristics of the municipalities included in the sample. Participants to the experiment are asked to play a two-round risk game. The first round consists of choosing among six gambles with the following payoffs:

Gamble choice	Low payoff (yellow)	High payoff (blue)	Expected value	Standard deviation	Risk aversion class	CRRA range
Gamble 1	3,000	3,000	3,000	0	Extreme	Infinity to 7.49
Gamble 2	2,700	5,700	4,200	2,121	Severe	7.49 to 1.73
Gamble 3	2,400	7,200	4,800	3,394	Intermediate	1.73 to 0.81
Gamble 4	1,800	9,000	5,400	5,091	Moderate	0.81 to 0.46
Gamble 5	1,000	11,000	6,000	7,071	Slight-neutral	0.46 to 0.00
Gamble 6	0	12,000	6,000	8,485	Neutral-negative	0 to −ve infinity

Source: Attanasio et al. (2012).

Figure 2.1: Gamble choices

In the second round, players can form risk-sharing groups and play the same game but share their payoffs among their group. As a relevant addition, players can privately renege from the risk-sharing group after observing their payoff. The risk-sharing game involves two levels of cooperation: the ex-ante willingness to join a risk-sharing group and the ex-post likelihood of reneging on the group. Individual characteristics and municipality characteristics can influence both levels of cooperation.

Given the difficulty in disentangling trust vis-à-vis social costs, the following information is used to proxy for different cooperative motives. The results of a public goods (PG) game played in the same laboratory experiment are used to proxy trust, while network ties are used to proxy social costs.

Indeed, as anticipated in the introduction, social costs are assumed to be stronger within networks due to repeated interactions. On the other hand, PG games are played with randomly selected groups. Hence, the exogeneity of group formation and the game's incentives help capturing cooperative atti-

tudes that are not strictly linked to informal costs, i.e., trust. Here follows a detailed description of the PG game and its incentives scheme.

The PG game consists of groups formed by around 40 players per municipality, who are asked to simultaneously choose whether to allocate a token valued at 2,000 Colombian pesos in the private or public account. In the latter case, they have to share the value with all the other players in the municipality. When contributing to the public account, the token is valued at double its amount. The game is played twice before playing the gambling game described earlier.

In each of the 70 municipalities selected for the experiment, 40 residents are randomly selected and invited to attend. However, some people did not show up when the experiment was run, and in some municipalities less than 40 people participated in the experiment (see Figure 2.2 in Appendix 2.7). The PG game is played among those who showed up in each municipality.

Contributions to the first round of the PG game are a function of network ties and trust (Evans and Krueger (2009)) and depend on the expected number of players cooperating n^E (Kurzban and Houser (2011)). However, for any given n^E, the expected return from cooperation is lower than the expected return from non-cooperation:

$$E[public_{account}] = \frac{4000}{40} + n^E \frac{4000}{40} < E[private_{account}] = 2000 + n^E \frac{4000}{40}$$

Given the game's incentives, people contributing in both rounds are assumed to be facing substantial costs from betraying any randomly formed group. Following Kurzban and Houser (2011) approach, players are divided into three groups according to their behaviour in the PG game: people contributing in both rounds are defined as a group of *cooperative* people. In

contrast, the others are considered *individualistic* if they never contributed or *mixed* if they changed choices across rounds. This variable is used in the empirical strategy as a proxy for trustworthiness, with *cooperative* individuals being considered more trustworthy than those classified as *mixed* or *individualistic*.

The main addition of this study to the existing literature on risk-sharing concerns the inclusion of municipality-level social cohesion as an explanatory variable for risk-sharing group formation. More specifically, it is assumed that when it comes to unknown individuals (i.e. people outside one's network), people might still find it profitable to engage in risk-sharing groups if they exhibit high levels of trust in the society. Following this reasoning, the share of individuals contributing to the second round of the PG game is used as a measure of social cohesiveness in each municipality and a measure of the level of trust towards the other people living in the municipality.

Tables 2.1 and 2.2 show some summary statistics on the characteristics of the participants in the laboratory experiment. Participants are primarily women (87%) aged 16 to 84 years old. More than 75% is married and about the same share lives in rural areas. When asked to play the risk-sharing game, on average, participants show a preference for gambles 3 and 4 (moderate/intermediate risk), and there is no significant difference across rounds. In round 2, only 8% of players renege after observing their payoff, 2% of whom after winning the lowest payoff.

The experimental data show that matching with people having different risk attitudes or outside the network of family and friends does not affect the probability of reneging[1] (see Figure 2.4 in Appendix 2.7). At the same time, looking at the overall distribution of players joining a risk-sharing group (85% of the total number of players), 891 players (45%) join groups of players having,

[1]For both sets of groups, the statistical tests cannot exclude the null hypothesis of equal means.

on average, higher risk gamble preferences, and 896 players (45%) group with players having on average lower risk gamble preferences (see Figure 2.4 in Appendix 2.7). Only 210 (10%) group with players with the same gamble preferences. Hence, participants form groups both within and outside close networks, in a way that leaves little space for reneging.

Looking at the PG game results, Table 2.3 shows that, in each municipality, on average 37% of players contribute to the first round of the PG game, 44% if we look at round 2. Moreover, the average share of family and friends ties, is very low, 1% for family and 7% for friends, while it is predominant the share of unknown people in the game (87%). A total of 25% of players contribute in both rounds, while the majority of players (45%) never contribute (see Figure 2.5 in Appendix 2.7). Lastly, 30% of people switch behaviour across rounds. As a result, the incentives to contribute in the second round are, on average, very low in the selected municipalities, and people choosing to contribute despite first round payoff are assumed to face high free-riding costs.

The next section describes the theoretical model used to study the steps of the decision-making process, which, through backward induction, elicits the incentives and disincentives underlying individuals' choice to join a risk-sharing group. The model has been simplified to a two-player setting, while the other game characteristics are modelled as in the experimental setting.

2.3 Nash equilibria in risk-sharing

Suppose two players must decide whether to play a gambling game individually or form a risk-sharing group and share the payoffs. Players are risk averse, and their utility function is modelled according to CARA: $U(x) = 1 - e^{-\alpha x}$.

In round 1, agents indicate their gamble preferences. There are two types

Table 2.1: Characteristics of participants

	mean	median	min	max	s.d.
Demographic characteristics					
Female	0.87	1	0	1	0.33
Age	41.72	40	16	84	11.37
Education = years of schooling	3.7	3	0	17	3.13
HH heads	0.29	0	0	1	0.45
Married	0.78	1	0	1	0.42
Live in municipal centre	0.34	0	0	1	0.48
Network characteristics					
Number of family ties recognised	0.32	0	0	5	0.67
Number of friendship ties recognised	2.42	2	0	16	2.57
Observations	2321				

Table 2.2: Gambling game results

	mean	median	min	max	s.d.
Won gamble in round 1	0.55	1	0	1	0.5
Joined a group	0.87	1	0	1	0.34
No. co-group members	3.69	3	0	28	3.9
Won gamble in round 2	0.58	1	0	1	0.5
Opted out of group	0.08	0	0	1	.27
Opted out of group having won in round 2	0.06	0	0	1	0.25
Gamble choice in round 1	3.61	4	1	6	1.5
Gamble choice in round 2	3.91	4	1	6	1.45
Observations	2321				

Table 2.3: Characteristics of PG game

	mean	median	min	max	standard deviation
Contribution 1st round	0.37	0	0	1	0.48
Contribution 2nd round	0.44	0	0	1	0.5
N. players	35.93	40	11	40	6.1
Share of family members	0.01	0	0	0.19	0.02
Share of friends	0.07	0.05	0	0.5	0.07
Share of acquaintances	0.15	.1	0	0.92	0.17
Observations	2321				

of gamble available:

- $a = \begin{cases} Y & \pi = \frac{1}{2} \\ y & \pi = \frac{1}{2} \end{cases}$ with expected value $x = \frac{Y}{2} + \frac{y}{2}$.
- $b = \begin{cases} Z & \pi = \frac{1}{2} \\ z & \pi = \frac{1}{2} \end{cases}$ with expected payoff $m = \frac{Z}{2} + \frac{z}{2}$.

Compared to gamble a, gamble b can be both riskier ($Z > Y$ and $z < y$) or less risky ($Z < Y$ and $z > y$), with $m > x$ if riskier and $m < x$ if less risky.

Definition 1: *Two players i and j are matched assortatively if both prefer gamble a. Instead, they are matched non-assortatively if i prefers gamble a and j prefers gamble b.*

When part of a risk-sharing group, after observing their gamble payoff, players can privately renege on the group. If players choose to renege, they only get their payoff, and the same holds for the other player.

Assumptions 1 and 2 regulate behaviour at indifference. Assumption 1 states that non-reneging is convenient only for strict dominance. This assumption is motivated by a conservative approach that aims to find the lowest cost such that non-reneging is optimal. Hence, results would hold also in case the assumption is relaxed. Similarly, Assumption 2 states that people only refuse a match when there is a strict disutility from it. This assumption is made for purely mathematical purposes, and its relaxation does not substantially affect the results.

Assumption 1: *In the absence of formal enforcement mechanisms, when indifferent, a player reneges.*

Assumption 2: *When indifferent between joining a risk-sharing group and playing alone, a player joins the group.*

To study the incentives to join a group (ex-ante) and to renege (ex-post), let's define two costs that add to the utility function, that each individual faces whenever:

- refuses to join a group – ex-ante cost $\equiv c \geq 0$
- reneges from the group – ex-post cost $\equiv r \geq 0$

The two costs can be interpreted in multiple ways. For instance, ex-ante costs c can be construed as peer pressure from joining a group, searching costs or even a strong preference for being part of a group rather than playing alone. Instead, ex-post costs r can be interpreted as altruism or preference for cooperation. Since reneging is a private decision, reputational costs are unlikely to arise in this setting.

2.3.1 Backwards induction: the decision to renege in t_2

Lemmas 1 and 2 describe the thresholds of r required for a player to strictly prefer pooling the payoff to reneging.

Lemma 1: *Under Assumption 1, when matched assortatively, reneging is always dominant for* $r \leq \frac{1}{\alpha} ln(1 + e^{\alpha x})$. *Instead, when* $r > \frac{1}{\alpha} ln(1 + e^{\alpha x})$, *non-reneging always dominates the option of reneging.*

Proof 1: see Appendix 2.6.

Lemma 2: *If matched non-assortatively, when* $r \leq \frac{1}{\alpha} ln(e^{\alpha(\frac{Y-Z}{2})} + e^{\alpha(\frac{Y-z}{2})})$ *it is always optimal to renege when playing with a higher-risk player. Instead, when playing with a lower-risk player, reneging is optimal only after observing the highest payoff. Again, when* $r > \frac{1}{\alpha} ln(e^{\alpha(\frac{Y-Z}{2})} + e^{\alpha(\frac{Y-z}{2})})$, *non-reneging always dominates the option of reneging.*

Proof 2: see Appendix 2.6.

Lemmas 1 and 2 show that when ex-post costs are non-binding, and players are matched in a non-assortative way, at least one of them has incentives to renege after observing the gamble payoff. The more so the less risk-averse players are.

For simplicity, the next section will focus only on the following case of non-assortativeness: $(Z > Y; z < y)$. The other option, $(Z < Y; z > y)$, is trivial.

2.3.2 Backwards induction: the decision to join a group in t_1

Lemma 3 describes individual preferences for group formation according to other players' risk preferences.

Lemma 3: *The more risk averse player i is, the more they prefer to group with a player $j \neq i$ having similar risk attitudes.*

Proof 3: see Appendix 2.6.

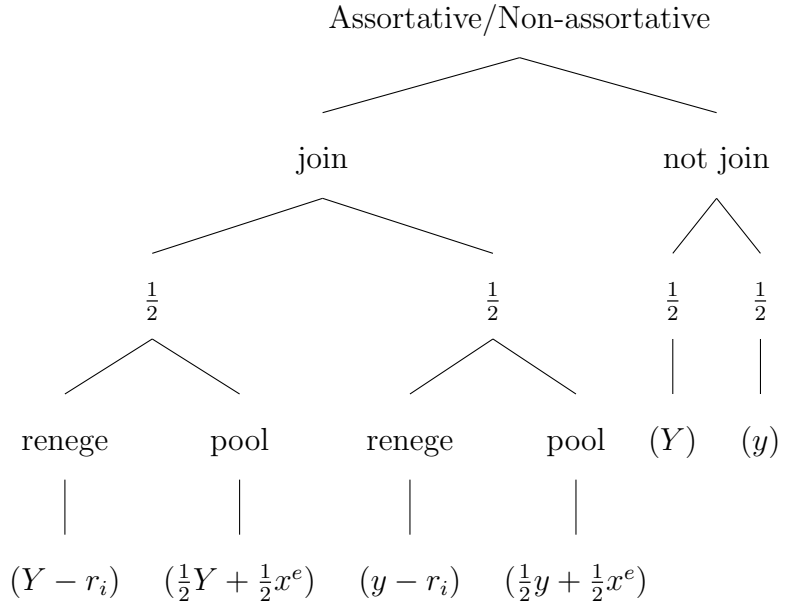
Lemma 3 describes how risk aversion affects the matching type (assortative or not). Indeed, risk-averse players are willing to match with lower-risk players more likely than they are willing to match with higher-risk players. However, for small differences in the risk of the gamble preferences, players would still find it optimal to share the risk rather than playing alone.

Following Lemmas 1-3, the first theoretical result is defined in Theorem 1, describing the incentives to join a risk-sharing group when player i faces high ex-post costs.

Theorem 1: *A player i facing binding ex-post costs from reneging (i.e. $r_i > \frac{1}{\alpha} ln(e^{\alpha(\frac{Y-Z}{2})} + e^{\alpha(\frac{Y-z}{2})})$ joins a risk-sharing group only with a player $j \neq i$ having similar reneging costs $r_j > \frac{1}{\alpha} ln(e^{\alpha(\frac{Y-Z}{2})} + e^{\alpha(\frac{Y-z}{2})})$.*

Proof : Let's summarise player i's decision-making process for assortative and non-assortative matching options.

The graph below summarises the decision-making process of player i. The beginning of the tree shows, for dyads of assortative and non-assortative players, the option of not joining or instead joining the group. In the latter case, the tree continues with the realisation of the individual gamble, after which the player faces the ex-post decision, in each node realised, of whether to renege or not. The expectations about the payoff that the player j will pool are indicated with x^e.



Since player i faces binding ex-post costs, in the left nodes (when joining the group), they will always pool the payoff. Hence, when deciding whether to join a group with player $j \neq i$ strictly depends on expectations about player

j's trustworthiness. The following steps will analyse the optimal choices in two cases:

1. player i expects player j to share the same reneging costs (i.e. according to Lemmas 1 and 2, both players will always pool the payoff)

2. player i expects player j to have different reneging costs (i.e. according to Lemmas 1 and 2, player j will pool the payoff only if strictly convenient)

Case 1: players i and j face the same reneging costs. Hence, for player i, the situation is equivalent to formal enforcement, and they can gain from sharing risk without risking that player j reneges and reduces player i's payoff. In this case, player i will join the group.

In the case of non-assortativeness, as shown in Appendix 2.6, the decision is also driven by the level of risk-aversion of player i. Hence, the optimal choice is not joining the group also whenever players exhibit extremely different levels of risk aversion.

Case 2: players i and j have different reneging costs. In this case, since player i is not going to renege ex-post, in the left nodes, player i would share the individual gamble results without the benefits of risk-sharing, since player j only pools the payoff when strictly convenient. Hence, player i has no incentives to join the group.

Intuitively, if player j is always expected to pool the payoff, for player i it is optimal to join the group. On the other hand, when expecting player j to be individualistic and hence pool only low payoffs, the losses generated from joining the group and pooling are certain. Hence, with no ex-ante costs, playing alone and not joining the group is the dominant option for player i.

Since decision-making is symmetric, the model predicts that when individuals face binding ex-post costs from reneging, they have incentives to join a

group together if the other player faces the similar ex-post costs. However, if they differ in ex-post costs, the player with high ex-post costs has no incentives to join the group, and the group will not be formed (Table 2.4).

Table 2.4: Theoretical predictions on group formation

Assortative		
j/i	r low	r high
r low	group	no group
r high	no group	group

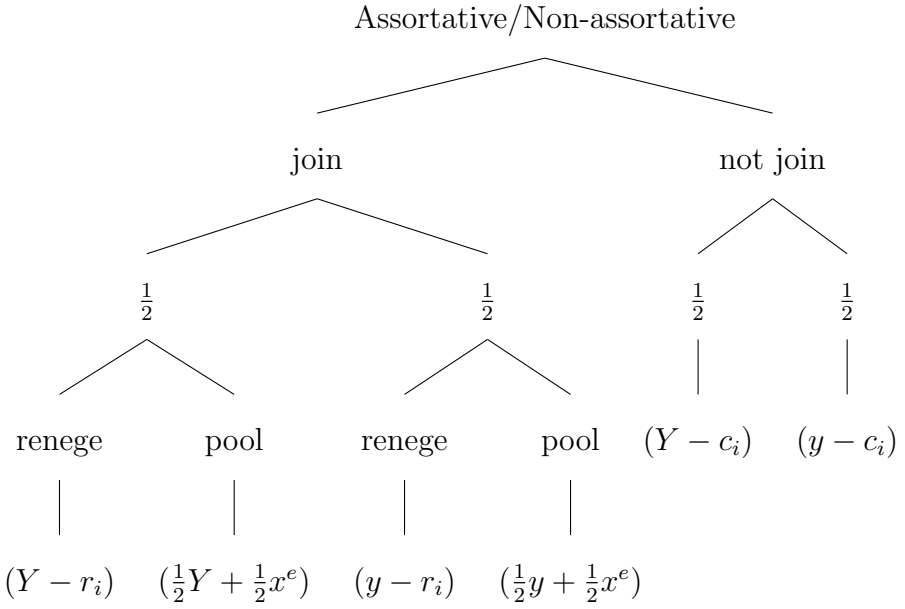
Non-assortative		
j/i	r low	r high
r low	group	no group
r high	no group	group

2.3.3 The role of ex-ante costs

Theorem 1 predicts that when trust is not reciprocal, trustworthy individuals have no incentives to join risk-sharing groups. However, ex-ante costs could compensate for the expected losses and incentivise trustworthy individuals to engage in risk-sharing groups even in absence of reciprocal trust. The second result of the model is defined in Theorem 2, which describes the level of ex-ante costs necessary to reverse the decision to join a group whenever ex-post costs are heterogeneous among agents.

Theorem 2: *In a setting without perfect enforcement, under Assumptions 1 and 2, ex-ante costs* $c \geq \frac{1}{\alpha}\, ln[\frac{e^{\alpha Y}+e^{\alpha y}}{e^{\alpha(\frac{Y+z}{2}\ )}+e^{\alpha(\frac{y+z}{2}\ )}}]$ *outweigh the disincentives to join a group when ex-post costs differ among agents.*

Proof : Consider **Case 2** from the proof of Theorem 1, and let's add to the option of not joining the group a social cost c.



For assortative matching, the level of c such that player i is willing to join the group is $E[U_i(join)] = \frac{1}{2}\ (1 - e^{-\alpha x}) + \frac{1}{2}\ (1 - e^{-\alpha y}) \geq \frac{1}{2}\ (1 - e^{-\alpha(Y-c)}) + \frac{1}{2}\ (1 - e^{-\alpha(y-c)}) = E[U_i(not\ join)] \implies c \geq \frac{1}{\alpha}\ ln[\frac{e^{\alpha Y} + e^{\alpha y}}{e^{\alpha x} + e^{\alpha y}}]$

On the other hand, for non-assortative matching, player i joins the group if $E[U_i(join)] = \frac{1}{2}\ (1 - e^{-\alpha(\frac{Y+z}{2}\ )}) + \frac{1}{2}\ (1 - e^{-\alpha(\frac{y+z}{2}\ )}) \geq \frac{1}{2}\ (1 - e^{-\alpha(Y-c)}) + \frac{1}{2}\ (1 - e^{-\alpha(y-c)}) = E[U_i(not\ join)] \implies c \geq \frac{1}{\alpha}\ ln[\frac{e^{\alpha Y} + e^{\alpha y}}{e^{\alpha(\frac{Y+z}{2}\ )} + e^{\alpha(\frac{y+z}{2}\ )}}]$

The more risk-averse a player is, the lower the threshold for c and the higher the likelihood of joining a group.

To conclude, the existence of ex-post costs that discourage reneging in t_2 can create disincentives to join a group in t_1, both assortatively and non-assortatively on risk preferences. On the other hand, ex-ante costs help overcome the negative effect by generating further incentives to join groups, guaranteeing group formation and ex-post cooperation from at least one agent.

The next section analyses players' behaviour in the Colombian laboratory experiment in light of the theoretical predictions. As anticipated in Section 2.2, people are divided into three groups according to their contributions to the PG games (cooperative, mixed and individualistic), and social trust is proxied

using a measure of social cohesion. The empirical analysis aims to study the individual propensity to join risk-sharing groups according to characteristics of trust, trustworthiness and the presence of social costs. To compare empirical results to theoretical predictions, three assumptions are made:

1. **Being *cooperative* means having high ex-post costs r.** Given the PG game incentives described in Section 2.2, people who cooperate in both rounds display a high cost for non-cooperation. Since the PG game is played with randomly chosen players, it is assumed that the same cost applies in the reneging decision, where groups are chosen endogenously.

2. **The stronger social cohesion is, the higher the likelihood that a player $j \neq i$ has high ex-post costs r_j.** A higher share of people contributing to the PG game is considered a symptom of an average high trustworthiness of people in the municipality. Hence, it is assumed that the likelihood that a player $j \neq i$ is trustworthy is higher in municipalities with stronger social cohesion.

3. **If two players i and j are part of the same network, both have high ex-ante costs c.** Due to repeated interactions, people can gain information about each other's preferences regarding cooperation and face pressure to join risk-sharing groups in the gambling game.

2.4 Empirical specification

The empirical specification relies on three pieces of information collected during the laboratory experiment described in Section 2.2. The first piece of information, the dependent variable, is measured using group formations in the risk-sharing game.

Instead, the second and third information, the two variables of interest, are (1) the network ties existing among participants, used to proxy for social costs, and (2) cooperative attitudes collected during the PG games, used to proxy for social incentives (trust and trustworthiness).

According to theoretical predictions, within networks (where ex-ante costs are high), one should observe a higher propensity to join risk-sharing groups than outside networks. Indeed, belonging to a network could balance off the effect of ex-post costs by increasing the ex-ante costs c. However, for fixed values of c, the expected returns from sharing risk are still higher if ex-post costs r are high for all the parties involved, and people are expected to show, even within networks, a preference for homogeneous matching in terms of r.

Outside networks, instead, where ex-ante costs are low due to the lack of repeated interactions, people are expected to match based on expected ex-post costs (r). In this case, social cohesion gives information on the average value of r in society. Hence, people are expected to show a higher propensity to form risk-sharing groups in more cohesive municipalities than in less cohesive ones.

To test the theoretical predictions, the empirical strategy follows three steps of specification:

1. The first step studies whether people are more likely to form risk-sharing groups within networks than outside networks.

2. The second step focuses on a within-networks analysis, to check whether reciprocal trust is still significant in explaining group formation within networks or, instead, being part of the same network is a sufficient condition for joining a group.

3. Lastly, the third step analyses whether people from more cohesive munic-

ipalities are more likely to form risk-sharing groups compared to people from less cohesive municipalities and how social cohesion affects the size of the groups formed.

In steps 1 and 2, a dyadic specification is used to study the likelihood of any possible pair of players being part of the same group. As described in detail by Fafchamps and Gubert (2007), the dyadic regression requires regressors to be symmetric. For this reason, all the regressors are expressed in terms of differences and combined levels of each dyad. Moreover, since dyadic observations are not independent, standard errors are clustered by pair. Instead, step 3 relies on a standard matrix specification built at the individual level and standard errors are clustered by municipality.

2.4.1 Step 1: network ties

The first step analyses the relationship between the likelihood that a dyad is part of the same risk-sharing group and network ties. The model estimated is the following:

$$(2.1)\ group_{ij} = \alpha + \beta_1 network_{ij} + \beta_2 cohesion_m + \beta_3 assortative_{ij} + \beta_5 X_{ij} + \epsilon$$

where

- $group_{ij} = \begin{cases} 0 & \text{if player } i \text{ and } j \text{ are not part of the same risk-sharing group} \\ 1 & \text{if player } i \text{ and } j \text{ are part of the same risk-sharing group} \end{cases}$
- $network_{ij} = \begin{cases} 0 & \text{if player } i \text{ and } j \text{ are not part of the same network} \\ 1 & \text{if player } i \text{ and } j \text{ are part of the same network} \end{cases}$
- $cohesion_m$ = share of players contributing to round 2 of the PG game in municipality m
- $assortative_{ij} = \begin{cases} 0 & \text{if player } i \text{ and } j \text{ have different risk preferences} \\ 1 & \text{if player } i \text{ and } j \text{ have the same risk preferences} \end{cases}$

- X_{ij} = sums and differences of players' characteristics (individual behaviour in the PG game, degree of family and friends in the session, first round gains in the gambling game, age, years of schooling, household consumption and household size)

As described in the theoretical model, differences in risk preferences are relevant in explaining the likelihood of forming a group. For this reason, the analysis includes the variable $assortative_{ij}$, in addition to the symmetric controls listed above (X_{ij}). Risk preferences are measured by the individual gamble chosen in the first round using the risk-sharing game.

Model (2.1) results are estimated using a probit regression. Coefficients are reported in Table 2.5. The coefficients on social cohesion and network ties are both positive and significant. In particular, the coefficient on network ties confirms the literature findings about stronger preferences for intra-network group formation. At the same time, controlling for network ties, the coefficient on social cohesion remains positive and significant, suggesting a higher likelihood for each dyad of joining the same risk-sharing group if part of a more cohesive municipality, no matter the presence of network ties between the pair.

Results highlight that networks play a significant role in boosting risk-sharing, which can be interpreted as an effect of social costs. However, controlling for network ties, social cohesion still plays a relevant role in explaining risk-sharing. Hence, social costs seem not to entirely drive the role of network ties in facilitating group formation. Instead, it is possible that the endogeneity of network formation leads to networks being homogeneous in terms of risk preferences, making it more profitable to join risk-sharing groups within networks rather than outside. Indeed, contrarily to theoretical predictions, risk preferences are not significant in the estimated model, probably homogeneity

being already captured by network ties. For this reason, step 2 focuses on the determinants of group formation within networks.

2.4.2 Step 2: reciprocal trust

In step 2, to explore the relevant characteristics in forming a risk-sharing group within networks, a linear probability model is used to estimate the following empirical model, selecting only dyads that are part of the same network[2].

$$(2.2)\ group_{ij} = \alpha + \beta_1(r_i \times r_j) + \beta_2 assortative_{ij} + \beta_3 X_{ij} + \beta 4 FE_m + \epsilon$$

where

- $group_{ij} = \begin{cases} 0 & \text{if player } i \text{ and } j \text{ are not part of the same risk-sharing group} \\ 1 & \text{if player } i \text{ and } j \text{ are part of the same risk-sharing group} \end{cases}$
- $r_{i,j} = \begin{cases} 1 = \text{individualistic} & \text{i.e. plays private in both PG game rounds} \\ 2 = \text{mixed} & \text{i.e. changes choice between PG game rounds} \\ 3 = \text{cooperative} & \text{i.e. plays public in both PG game rounds} \end{cases}$
- $assortative_{ij} = \begin{cases} 0 & \text{if player } i \text{ and } j \text{ have different risk preferences} \\ 1 & \text{if player } i \text{ and } j \text{ have the same risk preferences} \end{cases}$
- X_{ij} = sums and differences of players' characteristics (degree of family and friends in the session, gains in the first round of the gambling game, age, years of schooling, household consumption and household size)
- FE_m = municipality fixed effects.

As within the network it is expected that people have a better knowledge about each other's trustworthiness (Attanasio et al. (2012)), the explanatory variables include individual's pairwise responses to the PG games $(r_i \times r_j)$

[2]Four municipalities are excluded from the analysis because of the lack of variation in the dependent variable.

instead of the overall share of people contributing to the PG game introduced in equation (2.2) (i.e. variable $cohesion_v$).

In particular, within networks, under the assumption that risk-sharing group formation is motivated exclusively by social costs, coefficients on individuals' trustworthiness and cooperative attitudes should not be significant in explaining the likelihood of being in the same risk-sharing group, and the presence of network ties should be a sufficient condition to form a risk-sharing group.

Results are shown in Table 2.6. Coefficients on $(r_i \times r_j)$ are reported in reference to the category *both players being individualistic*. As predicted by the theoretical model, players having different cooperative attitudes (individualistic and mixed, cooperative and individualistic, etc.) display a lower likelihood of being in the same group compared to players being both individualistic. While being *both cooperative* does not significantly change the relative likelihood of forming a risk-sharing group. Again, contrarily to theoretical predictions, differences in risk preferences do not seem relevant in explaining variations in the probability of a dyad being part of the same risk-sharing group.

Results suggest that even in the presence of social costs (i.e. within networks), homogeneity in ex-post costs is a relevant factor in group formation. This is especially true if homogeneity concerns reciprocal trust (high r), although with a low statistical significance.

2.4.3 Step 3: social cohesion

The last step of the analysis involves a standard matrix with observations at the individual level to estimate the relationship between municipality-level social cohesion and two dependent variables: (1) the individual likelihood of

forming a risk-sharing group versus playing the gambling game as a single-player and (2) the size of the groups formed. In both equations, network links are added as controls since, as shown in the previous sections, they directly affect the willingness to form a risk-sharing group.

To estimate how social cohesion and network ties correlate with the likelihood of forming any risk-sharing group, the following model is estimated using a probit regression:

$$(2.3)\ join_i = \alpha + \beta_1 cohesion_m + \beta_2 ties_i + \beta_3 X_i + \epsilon$$

where,

- $join_i = \begin{cases} 0 & \text{if player } i \text{ chooses to play the gamble as a single-player} \\ 1 & \text{if player } i \text{ joins a risk-sharing group} \end{cases}$
- $cohesion_m$ = share of players contributing to round 2 of the PG game in municipality m
- $ties_i$ = number of family or friendship links recognised by player i
- X_i = characteristics of player i (individual behaviour in the PG game, risk preferences, age, years of schooling and first round gains in the gambling game)

The estimated coefficients are reported in Table 2.7, and show that social cohesion correlates positively and significantly to the likelihood of joining a risk-sharing group. Moreover, controlling for the level of social cohesion in each municipality, the number of network ties recognised by each player becomes non statistically significant in the analysis. Hence, coming from a more cohesive municipality, on average, increases the likelihood of joining a risk-sharing group, while being part of a larger network seems not to affect the likelihood of joining a risk-sharing group compared to individuals who are part of smaller networks.

Lastly, to estimate how social cohesion and network ties are correlated with the size of the risk-sharing groups formed, the following model is estimated using a linear regression:

$$(2.4)\ size_g = \alpha + \beta_1 cohesion_m + \beta_2 network_{degree} + \beta_3 n_{players} + \epsilon$$

where,

- $size_g$ = group size
- $cohesion_m$ = share of players contributing to round 2 of the PG game in municipality m
- $network_{degree}$ = average degree of friendship and family links in the session
- $n_{players}$ = total number of players in the experiment.

Table 2.8 reports the estimation results of model (2.4). Coefficients show that municipalities with higher social cohesion also have, on average, groups formed by more members. At the same time, the share of network ties in the session correlates negatively with the group size, suggesting that stronger network links may result in smaller groups. However, although it is relatively large in magnitude, the coefficient is not statistically significant.

Overall, results confirm that higher social cohesion increases trust in others, regardless of individual attitudes towards cooperation. This is most critical outside networks, where the lack of information might discourage group formation. Moreover, the absence of network ties increases the likelihood of uncorrelated incomes. Hence, risk-sharing groups formed on the basis of social cohesion are potentially also more efficient (Fafchamps and Gubert (2007)).

Results on sizes are relevant in the context of efficiency, too. Indeed, the bigger the groups formed, the lower the risk faced by its members and the higher the gains from sharing risk. However, bigger groups could increase the

incentives to renege by lowering the chances of being caught reneging. This might be among the reasons why we observe smaller groups when social costs (i.e. network ties) enter as a driver of group formation.

	(1)	(2)
VARIABLES	group	group
network	0.737***	0.744***
	(0.018)	(0.018)
cohesion		0.380***
		(0.022)
assortative	-0.003	-0.008
	(0.014)	(0.014)
Constant	-1.707***	-2.408***
	(0.202)	(0.221)
Controls	yes	yes
FE	no	no
Observations	67,398	67,398

Table 2.5: Probability of being part of the same group

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The table shows equation (2.1) coefficients. Coefficients are estimated with a probit regression on the dyadic matrix to explain variations occurring in the likelihood that each pair of players is part of the same risk-sharing group. The variable *network* is equal to 1 if the pair of players is part of the same network of family and friends and 0 otherwise. The variable *cohesion* indicates the share of individuals in each municipality that contributed to the second round of the PG game. The variable *assortative* is equal to 1 if the pair of players has the same risk preferences and 0 otherwise. Risk preferences are measured by the individual gamble chosen in the first step of the risk-sharing game. The controls include individual cooperative attitudes measured in the PG game, degree of family and friends in the session, gains in the first round of the gambling game, age, years of schooling, household consumption and household size. Standard errors are bootstrapped with 1,000 replications and clustered by pair. Coefficients on the other covariates are reported in the Appendix 2.7.

VARIABLES	(1) group
both cooperative	0.049**
	(0.023)
cooperative and individualistic	-0.120***
	(0.023)
cooperative and mixed	-0.075***
	(0.023)
individualistic and mixed	-0.068***
	(0.018)
both mixed	0.022
	(0.021)
assortative	0.007
	(0.009)
Constant	0.486***
	(0.164)
Controls	yes
FE	municipality
Observations	7,632
R-squared	0.152

Table 2.6: Probability of being part of the same group within networks

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The table shows equation (2.2) coefficients. Coefficients are estimated with a linear probability model on the dyadic matrix to explain variations occurring in the likelihood that each pair of players is part of the same risk-sharing group. The analysis is carried out only for dyads part of the same network of family and friends. The coefficients on *both cooperative*, *cooperative and individualistic*, *cooperative and mixes*, *individualistic and mixed* are reported relative to players being *both individualistic*, and indicate the individual attitudes towards cooperation shown in the PG games. The variable *assortative* is equal to 1 if the pair of players has the same risk preferences and 0 otherwise. Risk preferences are measured by the individual gamble chosen in the first step of the risk-sharing game. The controls include degree of family and friends in the session, gains in the first round of the gambling game, age, years of schooling, household consumption and household size. Standard errors are bootstrapped with 1,000 replications and clustered by pair. Coefficients on the other covariates are reported in the Appendix 2.7.

VARIABLES	(1) join
cohesion	1.093**
	(0.433)
ties	0.045
	(0.037)
Constant	0.718**
	(0.297)
Controls	yes
FE	no
Observations	2,112

Table 2.7: Probability of joining any group

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The table shows equation (2.3) coefficients. Coefficients are estimated with a probit regression to explain variations occurring in the individual likelihood of forming a risk-sharing group. The variable *cohesion* indicates the share of individuals in each municipality that contributed to the second round of the PG game. The variable *ties* indicates the number of network ties recognised by player i. The controls include the individual cooperative attitudes measured in the PG game, gamble preference, age, years of schooling and gains in the first round of the gambling game. Standard errors are clustered at the municipal level. Coefficients on the other covariates are reported in the Appendix 2.7.

VARIABLES	(1) size
cohesion	1.862***
	(0.589)
network degree	-2.122
	(2.580)
n players	0.009
	(0.021)
Constant	1.844**
	(0.733)
Controls	no
FE	no
Observations	849
R-squared	0.078

Table 2.8: Determinants of group size

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.The table shows equation (2.4) coefficients. Coefficients are estimated with a linear regression model to explain variations occurring in the size of the risk-sharing groups formed. The variable *cohesion* indicates the share of individuals in each municipality that contributed to the second round of the PG game. The variable *network degree* indicates the average degree of network links in the experimental session. The variable *n players* counts the number of players that attended the laboratory experiment. Standard errors are clustered at the municipal level.

2.5 Conclusions

The laboratory experiment in Colombia collects relevant evidence on how people differ in cooperative attitudes and engage in informal risk-sharing contracts. This chapter explores the individual incentives underlying the decision to share risk, focusing on the effect of trust and social cohesion versus social costs.

The analysis first includes a simple theoretical framework defining the optimal choices in a risk-sharing setting and its decision thresholds. Then, the theoretical predictions are tested in an empirical framework, using experimental data collected in 70 Colombian municipalities. The experimental framework guarantees the isolation of the mechanisms involved in risk-sharing.

The study confirms previous literature findings about people preference to join risk-sharing groups intra-networks. Indeed, the existence of network ties is correlated with a higher likelihood of forming a risk-sharing group. However, within and outside networks, reciprocal trust strongly correlates with group formation, suggesting that it is a paramount factor in facilitating risk-sharing group formation, even in the presence of social costs.

As an addition to the existing literature on risk-sharing, this chapter highlights the role of social cohesion in forming risk-sharing groups. The economic literature has mainly focused on risk-sharing agreements within networks, ignoring how this practice can be incentivised outside networks. This work finds that social cohesion helps overcome the individual disincentives to cooperation by increasing the overall trust in society. This is relevant since it also involves unknown individuals, showing that it is possible to expand risk sharing to individuals whose income is uncorrelated (Fafchamps and Gubert (2007)). Moreover, the groups formed under higher social cohesion tend to be larger, increasing the benefits from sharing risk.

Results have significant policy implications, as they suggest boosting cooperation among individuals under minimal conditions that increase social cohesiveness. For instance, as shown by Tajfel (1970) and Akerlof and Kranton (2000), one can improve expectations about others and reduce hostility towards unknown people by creating a sense of identity or belonging. Indeed, as Charness, Cobo-Reyes, and Jimenez (2014) highlight, the mere introduction of team building activities can significantly increase cooperation and reduce discrimination among people that are part of different groups.

Results rely on a lab-in-the-field experiment designed to investigate group formation in risk-sharing. Hence, it requires validation also in natural settings.

2.6 Appendix A

Theoretical proofs

Proof 1: player i faces the following decision

1. i observes Y and reneges if $(1-e^{-\alpha(Y-r)}) \geq \frac{1}{2}\left(1-e^{-\alpha Y}\right)+\frac{1}{2}\left(1-e^{-\alpha x}\right)$

   $\implies 1-e^{-\alpha(Y-r)} \geq 1-e^{-\alpha Y}-e^{-\alpha x}$

   $\implies e^{-\alpha Y}-e^{-\alpha Y}e^{\alpha r} \geq -e^{-\alpha x}$

   $\implies 1+e^{\alpha x} \geq e^{\alpha r}$

   $\implies r \leq \frac{1}{\alpha}\ \ln(1+e^{\alpha x})$.

   This condition holds both if a player always expects the other to pool the payoff and expects to pool only the lowest payoff.

2. i observes y and reneges if

   (a) expects the other player to always pool the payoff: $(1-e^{-\alpha(y-r)}) \geq \frac{1}{2}\left(1-e^{-\alpha y}\right)+\frac{1}{2}\left(1-e^{-\alpha x}\right)$, which is never true since $y<x$;

   (b) expects the other player to pool only the lowest payoff: $(1-e^{-\alpha(y-r)}) \geq (1-e^{-\alpha y})$ if $r=0$ player i is indifferent.

To conclude, when matched assortatively, player i reneges only when winning the highest payoff.

Proof 2: Suppose player i is less risky and thinks player j is cooperative (and pools every payoff).

1. player i observes Y and reneges if

   $(1-e^{-\alpha(Y-r)}) \geq \frac{1}{2}\left(1-e^{-\alpha\left(\frac{Y+Z}{2}\right)}\right)+\frac{1}{2}\left(1-e^{-\alpha\left(\frac{Y+z}{2}\right)}\right)$

   $\implies 1-e^{-\alpha Y}e^{\alpha r} \geq 1-e^{-\alpha\left(\frac{Y+Z}{2}\right)}-e^{-\alpha\left(\frac{Y+z}{2}\right)}$

   $\implies e^{\alpha r} \leq e^{\alpha\left(\frac{Y-Z}{2}\right)}+e^{-\alpha\left(\frac{Y-z}{2}\right)}$

   $\implies r \leq \frac{1}{\alpha}\ \ln(e^{\alpha\left(\frac{Y-Z}{2}\right)}+e^{\alpha\left(\frac{Y-z}{2}\right)})$.

2. player i observes y, instead, and reneges if $r \leq \frac{1}{\alpha} \ln(e^{\alpha(\frac{y-Z}{2})} + e^{\alpha(\frac{y-z}{2})})$.

On the other hand, if player j is individualistic, when observing Z, since $Z > Y > y$, player j will renege, and, when observing z, since $z < y < Y$, will pool the payoff. Hence, if player i thinks player j is individualistic:

1. if player i observes Y, it follows from Proof 1 that they will renege if $r \leq \frac{1}{\alpha} \ln(1 + e^{\alpha(\frac{Y-z}{2})})$;
2. if player i observes y, instead, will renege if $r \leq \frac{1}{\alpha} \ln(1 + e^{\alpha(\frac{y-z}{2})})$.

The opposite holds if player i is riskier than player j. Hence, if $r \leq \frac{1}{\alpha} \ln(e^{\alpha(\frac{y-Z}{2})} + e^{\alpha(\frac{y-z}{2})})$, reneging is optimal for player i even if player j is assumed to be cooperative. However, if player j is less risky than player i, reneging when $r \leq \frac{1}{\alpha} \ln(e^{\alpha(\frac{y-Z}{2})} + e^{\alpha(\frac{y-z}{2})})$ is optimal only after observing the highest payoff.

Proof 3: If ex-post costs are binding for both players, player i will accept to group together with player j accordingly:

- When playing alone the expected utility is $EU_S = \frac{1}{2}\left(1 - e^{-\alpha Y}\right) + \frac{1}{2}\left(1 - e^{-\alpha y}\right)$
- Instead, when playing with j, the expected utility will be
  $U(win_i|win_j) = 1 - e^{-\alpha\left(\frac{Y}{2} + \frac{Z}{2}\right)}$
  $U(win_i|loose_j) = 1 - e^{-\alpha\left(\frac{Y}{2} + \frac{z}{2}\right)}$
  $U(loose_i|win_j) = 1 - e^{-\alpha\left(\frac{Z}{2} + \frac{y}{2}\right)}$
  $U(loose_i|loose_j) = 1 - e^{-\alpha\left(\frac{y}{2} + \frac{z}{2}\right)}$

Assuming that players will never renege, then player i will accept to group with player j if:

$$EU_S \leq EU_G \implies \tfrac{1}{2}\left(1-e^{-\alpha Y}\right)+\tfrac{1}{2}\left(1-e^{-\alpha y}\right) \leq \tfrac{1}{4}\left(1-e^{-\alpha\left(\frac{Y}{2}+\frac{Z}{2}\right)}\right)+$$
$$\tfrac{1}{4}\left(1-e^{-\alpha\left(\frac{y}{2}+\frac{z}{2}\right)}\right)+\tfrac{1}{4}\left(1-e^{-\alpha\left(\frac{Y}{2}+\frac{z}{2}\right)}\right)+\tfrac{1}{4}\left(1-e^{-\alpha\left(\frac{Z}{2}+\frac{y}{2}\right)}\right)$$

$$\implies 2-e^{-\alpha Y}-e^{-\alpha y} \leq 2-\tfrac{1}{2}e^{-\alpha\left(\frac{Y}{2}+\frac{Z}{2}\right)}-\tfrac{1}{2}e^{-\alpha\left(\frac{y}{2}+\frac{z}{2}\right)}-\tfrac{1}{2}e^{-\alpha\left(\frac{Y}{2}+\frac{z}{2}\right)}-$$
$$\tfrac{1}{2}e^{-\alpha\left(\frac{y}{2}+\frac{Z}{2}\right)}$$

$$\implies -e^{-\alpha Y}-e^{-\alpha y} \leq -\tfrac{1}{2}e^{-\alpha\frac{Y}{2}}e^{-\alpha\frac{Z}{2}}-\tfrac{1}{2}e^{-\alpha\frac{Y}{2}}e^{-\alpha\frac{z}{2}}-\tfrac{1}{2}e^{-\alpha\frac{y}{2}}e^{-\alpha\frac{Z}{2}}-$$
$$\tfrac{1}{2}e^{-\alpha\frac{y}{2}}e^{-\alpha\frac{z}{2}}$$

$$\implies -e^{-\alpha Y}+\tfrac{1}{2}e^{-\alpha\frac{Y}{2}}e^{-\alpha\frac{Z}{2}}+\tfrac{1}{2}e^{-\alpha\frac{Y}{2}}e^{-\alpha\frac{z}{2}} \leq e^{-\alpha y}-\tfrac{1}{2}e^{-\alpha\frac{y}{2}}e^{-\alpha\frac{Z}{2}}-\tfrac{1}{2}e^{-\alpha\frac{y}{2}}e^{-\alpha\frac{z}{2}}$$

$$\implies e^{-\alpha\frac{Y}{2}}\left(\tfrac{1}{2}e^{-\alpha\frac{Z}{2}}+\tfrac{1}{2}e^{-\alpha\frac{z}{2}}-e^{-\alpha\frac{Y}{2}}\right) \leq e^{-\alpha\frac{y}{2}}\left(e^{-\alpha\frac{y}{2}}-\tfrac{1}{2}e^{-\alpha\frac{Z}{2}}-\tfrac{1}{2}e^{-\alpha\frac{z}{2}}\right)$$

Since $Y > y$, $e^{-\alpha\frac{Y}{2}} < e^{-\alpha\frac{y}{2}}$ is satisfied. Hence, let's look at the conditions under which $\frac{1}{2}e^{-\alpha\frac{Z}{2}}+\frac{1}{2}e^{-\alpha\frac{z}{2}}-e^{-\alpha\frac{Y}{2}} \leq e^{-\alpha\frac{y}{2}}-\frac{1}{2}e^{-\alpha\frac{Z}{2}}-\frac{1}{2}e^{-\alpha\frac{z}{2}}$ is also satisfied. Simplifying the inequality, this means:

$$\implies e^{-\alpha\frac{Z}{2}}+e^{-\alpha\frac{z}{2}} \leq e^{-\alpha\frac{Y}{2}}+e^{-\alpha\frac{y}{2}}$$

$$\implies e^{\alpha\frac{Z}{2}}-e^{\alpha\frac{Y}{2}} \geq e^{\alpha\frac{y}{2}}-e^{\alpha\frac{z}{2}}$$

which is true if

1. For $e^{\alpha\frac{z}{2}} = e^{\alpha\frac{y}{2}} \implies e^{\alpha\frac{Z}{2}} \geq e^{\alpha\frac{Y}{2}}$. This case can only happen in the empirical setting if the two gambles are equal.
2. For $e^{\alpha\frac{z}{2}} > e^{\alpha\frac{y}{2}} \implies e^{\alpha\frac{Z}{2}}-e^{\alpha\frac{Y}{2}} \geq -\epsilon$, with $\epsilon = \left(e^{\alpha\frac{y}{2}}-e^{\alpha\frac{z}{2}}\right) > 0$. This case is always satisfied for $e^{\alpha\frac{Z}{2}} \geq e^{\alpha\frac{Y}{2}}$ (and $Z \geq Y$), not always for $e^{\alpha\frac{Z}{2}} < e^{\alpha\frac{Y}{2}}$ (and $Z < Y$).
3. For $e^{\alpha\frac{z}{2}} < e^{\alpha\frac{y}{2}} \implies e^{\alpha\frac{Z}{2}}-e^{\alpha\frac{Y}{2}} \geq \epsilon$. This case is never satisfied for $Z \leq Y$ but can be satisfied for some values of $Z > Y$.

To conclude, player i will always group with the same risk players (case

1). Instead, it is unclear whether they are willing group with a lower-risk player or a riskier one. In case of more risk, player i will group only if the difference in the utility gains for the higher payoffs compensates for the differences in the utility losses for the lower payoffs. The opposite is true for less risky gambles. Since decision-making is symmetrical, the more risk-averse the players are, the more likely they are to group with players with the same gamble preferences. Instead, if risk aversion is low, they are willing to group together even if they differ in gamble preferences.

Let's check how this applies in the context of the lab experiment with a few examples.

- Case 1: player i has a preference for gamble 3 and has to decide whether to match with the following other players.
  - player j prefers gamble 2:
    $e^{-\alpha\frac{7200}{2}}\left(\frac{1}{2}e^{-\alpha\frac{5700}{2}}+\frac{1}{2}e^{-\alpha\frac{2700}{2}}-e^{-\alpha\frac{7200}{2}}\right)\leq$
    $e^{-\alpha\frac{2400}{2}}\left(e^{-\alpha\frac{2400}{2}}-\frac{1}{2}e^{-\alpha\frac{5700}{2}}-\frac{1}{2}e^{-\alpha\frac{2700}{2}}\right)$
    $\Longrightarrow \frac{1}{2}e^{-\alpha 4500}+e^{-\alpha 2550}-e^{-\alpha 7200}+\frac{1}{2}e^{-\alpha 150}\leq 1$,which is true only for $\alpha\in[-0.0003;0]$
  - player j prefers gamble 4:
    $e^{-\alpha\frac{7200}{2}}\left(\frac{1}{2}e^{-\alpha\frac{9000}{2}}+\frac{1}{2}e^{-\alpha\frac{1800}{2}}-e^{-\alpha\frac{7200}{2}}\right)\leq$
    $e^{-\alpha\frac{2400}{2}}\left(e^{-\alpha\frac{2400}{2}}-\frac{1}{2}e^{-\alpha\frac{9000}{2}}-\frac{1}{2}e^{-\alpha\frac{1800}{2}}\right)$
    $\Longrightarrow -\frac{1}{2}e^{-\alpha 8100}-\frac{1}{2}e^{-\alpha 4500}+e^{-\alpha 7200}+e^{-\alpha 300}-\frac{1}{2}e^{-\alpha 5700}\geq\frac{1}{2}$,which is true for $\alpha\in[0;0.002]$
  - player j prefers gamble 5:
    $e^{-\alpha\frac{7200}{2}}\left(\frac{1}{2}e^{-\alpha\frac{11400}{2}}+\frac{1}{2}e^{-\alpha\frac{600}{2}}-e^{-\alpha\frac{7200}{2}}\right)\leq$
    $e^{-\alpha\frac{2400}{2}}\left(e^{-\alpha\frac{2400}{2}}-\frac{1}{2}e^{-\alpha\frac{11400}{2}}-\frac{1}{2}e^{-\alpha\frac{600}{2}}\right)$
    $\Longrightarrow \frac{1}{2}e^{-\alpha 9300}+\frac{1}{2}e^{-\alpha 3900}-e^{-\alpha 7200}\leq e^{-\alpha 2400}-\frac{1}{2}e^{-\alpha 6900}-$

$\frac{1}{2}e^{-\alpha 1500} \implies e^{-\alpha 5700} + e^{-\alpha 900} - \frac{1}{2}e^{-\alpha 7800} - \frac{1}{2}e^{-\alpha 2400} - \frac{1}{2}e^{-\alpha 5400} \geq \frac{1}{2}$, which is only true for positive $\alpha \in [0; 0.0005]$

Hence, player i strongly prefers matching with same-risk players if risk-averse $\alpha > 0$.

- Case 2: player i prefers gamble 1
  - player j prefers gamble 6:
    $e^{-\alpha\frac{3000}{2}}\left(\frac{1}{2}e^{-\alpha\frac{12000}{2}} + \frac{1}{2} - e^{-\alpha\frac{3000}{2}}\right) \leq$
    $e^{-\alpha\frac{3000}{2}}\left(e^{-\alpha\frac{3000}{2}} - \frac{1}{2}e^{-\alpha\frac{12000}{2}} - \frac{1}{2}\right)$
    $\implies e^{-\alpha 7500} + e^{-\alpha 1500} - 2e^{-\alpha 3000} \leq 0$, which is true for $\alpha \in [0; 0.0004]$
  - player j prefers gamble 3:
    $e^{-\alpha\frac{3000}{2}}\left(\frac{1}{2}e^{-\alpha\frac{7200}{2}} + \frac{1}{2}e^{-\alpha\frac{2400}{2}} - e^{-\alpha\frac{3000}{2}}\right) \leq$
    $e^{-\alpha\frac{3000}{2}}\left(e^{-\alpha\frac{3000}{2}} - \frac{1}{2}e^{-\alpha\frac{7200}{2}} - \frac{1}{2}e^{-\alpha\frac{2400}{2}}\right)$
    $\implies e^{-\alpha 5100} + e^{-\alpha 2700} - 2e^{-\alpha 3000} \leq 0$, which is true for $\alpha \in [0; 0.0023]$
  - player j prefers gamble 2:
    $e^{-\alpha\frac{3000}{2}}\left(\frac{1}{2}e^{-\alpha\frac{5700}{2}} + \frac{1}{2}e^{-\alpha\frac{2700}{2}} - e^{-\alpha\frac{3000}{2}}\right) \leq$
    $e^{-\alpha\frac{3000}{2}}\left(e^{-\alpha\frac{3000}{2}} - \frac{1}{2}e^{-\alpha\frac{5700}{2}} - \frac{1}{2}e^{-\alpha\frac{2700}{2}}\right)$
    $\implies e^{-\alpha 4350} + e^{-\alpha 2850} - 2e^{-\alpha 3000} \leq 0$, which is true for $\alpha \in [0; 0.005]$

Player i, in this case, has a strong preference for same-risk grouping but is willing to accept to match with a median risk aversion player if α is low enough.

- Case 3: player i prefers gamble 6
  - player j prefers gamble 1:

$e^{-\alpha 7500} - e^{-\alpha 12000} \leq 1 - e^{-\alpha 1500} \implies e^{-\alpha 7500} - e^{-\alpha 12000} + e^{-\alpha 1500} \leq 1$, which is always true for $\alpha \geq 0.0001$ as well as for $\alpha \leq 0.2$

- player j prefers gamble 3:
  $e^{-\alpha \frac{12000}{2}} \left(\frac{1}{2} e^{-\alpha \frac{7200}{2}} + \frac{1}{2} e^{-\alpha \frac{2400}{2}} - e^{-\alpha \frac{12000}{2}} \right) \leq 1 - \frac{1}{2} e^{-\alpha \frac{7200}{2}} - \frac{1}{2} e^{-\alpha \frac{2400}{2}}$
  $\implies e^{-\alpha 9600} + e^{-\alpha 7200} - 2e^{-\alpha 12000} + e^{-\alpha 3600} + e^{-\alpha 1200} \leq 2$, which is true for $\alpha \geq 0.00005$ as well as for $\alpha \leq 0$

In this case, if player i is risk averse, he is always willing to match with less risky players.

- Case 4: player i prefers gamble 5
  - player j prefers gamble 4:
    $e^{-\alpha \frac{11400}{2}} \left(\frac{1}{2} e^{-\alpha \frac{9000}{2}} + \frac{1}{2} e^{-\alpha \frac{1800}{2}} - e^{-\alpha \frac{11400}{2}} \right) \leq$
    $e^{-\alpha \frac{600}{2}} \left(e^{-\alpha \frac{600}{2}} - \frac{1}{2} e^{-\alpha \frac{9000}{2}} - \frac{1}{2} e^{-\alpha \frac{1800}{2}} \right)$
    $\implies e^{-\alpha 9600} + e^{-\alpha 6000} + e^{-\alpha 4200} + e^{-\alpha 600} - 2e^{-\alpha 10800} \leq 2$, which is true for $\alpha \leq 0$ as well as for $\alpha \geq 0.00003$
  - player j prefers gamble 3:
    $e^{-\alpha \frac{11400}{2}} \left(\frac{1}{2} e^{-\alpha \frac{7200}{2}} + \frac{1}{2} e^{-\alpha \frac{2400}{2}} - e^{-\alpha \frac{11400}{2}} \right) \leq$
    $e^{-\alpha \frac{600}{2}} \left(e^{-\alpha \frac{600}{2}} - \frac{1}{2} e^{-\alpha \frac{7200}{2}} - \frac{1}{2} e^{-\alpha \frac{2400}{2}} \right)$
    $\implies e^{-\alpha 8700} + e^{-\alpha 6300} + e^{-\alpha 3300} + e^{-\alpha 900} - 2e^{-\alpha 10800} \leq 2$, which is true for $\alpha \leq 0$ as well as for $\alpha \geq 0.00006$
  - player j prefers gamble 6:
    $e^{-\alpha \frac{11400}{2}} \left(\frac{1}{2} e^{-\alpha \frac{12000}{2}} + \frac{1}{2} - e^{-\alpha \frac{11400}{2}} \right) \leq e^{-\alpha \frac{600}{2}} \left(e^{-\alpha \frac{600}{2}} - \frac{1}{2} e^{-\alpha \frac{12000}{2}} - \frac{1}{2} \right)$
    $\implies e^{-\alpha 11400} + e^{-\alpha 5400} + e^{-\alpha 6000} - 2e^{-\alpha 11100} - 2e^{-\alpha 300} \leq -1$,
    which is true for $\alpha \leq 0$ as well as for $\alpha \in [-0.002; 0.002]$

2.7 Appendix B

Figures

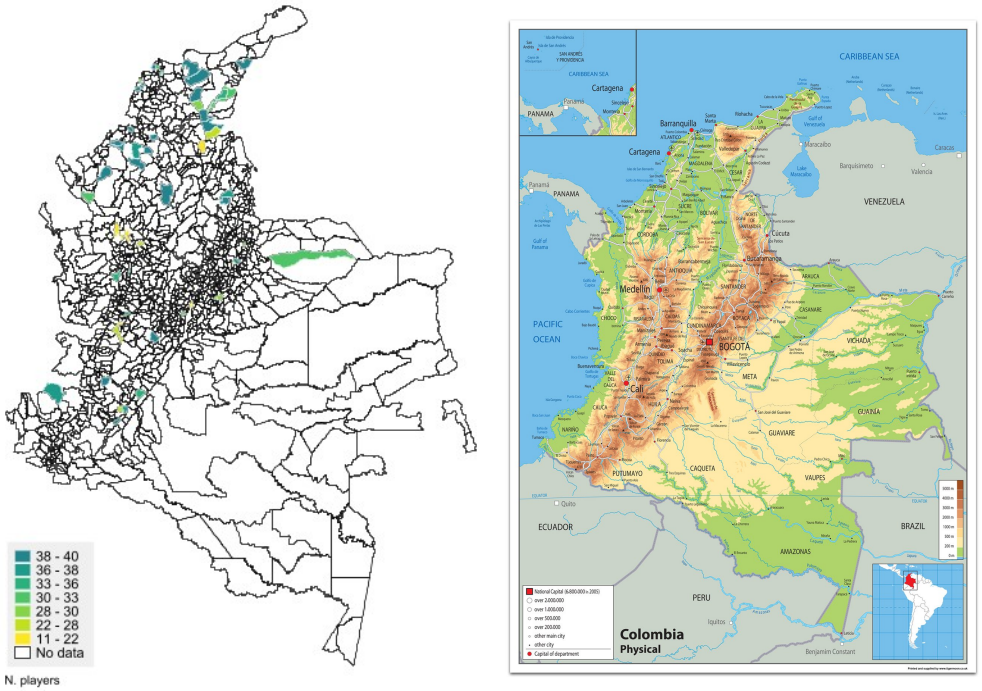


Figure 2.2: Municipalities selected and number of players

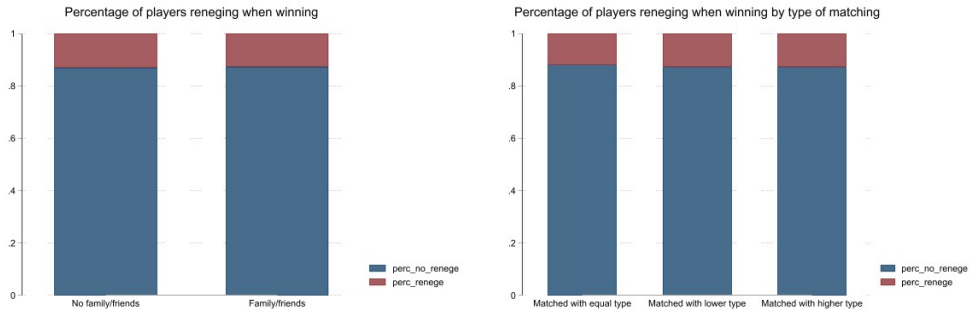


Figure 2.3: Percentage of players reneging

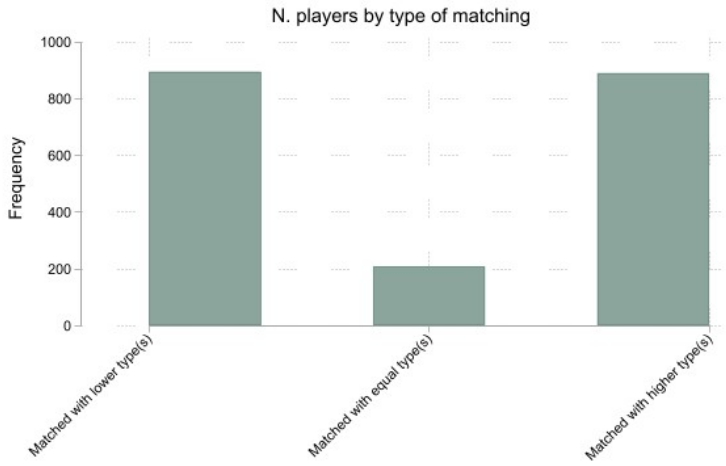


Figure 2.4: Type of matching

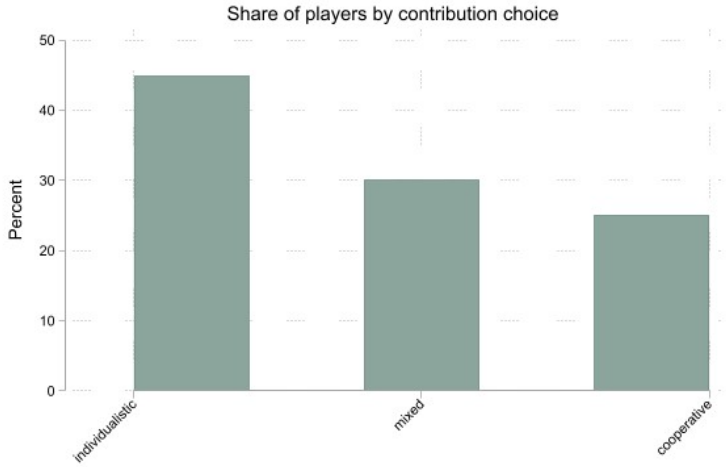


Figure 2.5: Share of players by contribution choice in the PG game

Regressions controls coefficients

VARIABLES	(1) group	(2) group
difference in PG game contributions	0.000	0.000
	(0.007)	(0.007)
sum of PG game contributions	-0.055***	-0.040***
	(0.006)	(0.006)
difference in round 1 winnings	-0.015***	-0.014***
	(0.002)	(0.002)
sum of round 1 winnings	-0.002**	-0.002
	(0.001)	(0.001)
difference in age	-0.003***	-0.003***
	(0.001)	(0.001)
sum of ages	-0.001*	-8.47e-05
	(0.001)	(0.001)
difference in years of schooling	-0.005**	-0.006**
	(0.003)	(0.003)
sum of years of schooling	0.004**	0.003*
	(0.002)	(0.002)
difference in log household consumption	-0.081***	-0.074***
	(0.015)	(0.015)
sum of log household consumption	0.034***	0.050***
	(0.008)	(0.009)
difference in hosuehold sizes	0.001	0.001
	(0.003)	(0.003)
sum of household sizes	-0.005***	-0.007***
	(0.002)	(0.002)
Observations	67,398	67,398

Table 2.9: Probability of being part of the same group: controls coefficients

VARIABLES	(1) group
difference in round 1 winnings	-0.009***
	(0.002)
sum of round 1 winnings	-0.0004
	(0.001)
difference in age	0.0001
	(0.001)
sum of ages	-0.001***
	(0.0003)
difference in years of schooling	-0.003*
	(0.002)
sum of years of schooling	0.002
	(0.001)
difference in log household consumption	-0.007
	(0.009)
sum of log household consumption	-0.002
	(0.007)
difference in hosuehold sizes	0.001
	(0.002)
sum of household sizes	-0.003**
	(0.001)
Observations	7,632

Table 2.10: Probability of being part of the same group within networks: controls coefficients

VARIABLES	(1) join
lottery choice in round 1	-0.036
	(0.023)
winnings from round 1	-7e-06
	(1e-05)
age	-0.002
	(0.004)
years of schooling	0.017
	(0.014)
Observations	2,112

Table 2.11: Probability of joining any group: controls coefficients

Chapter 3

Social cohesion and agriculture in Colombia

3.1 Introduction

Social cohesion is a very important factor for economic development since it affects individuals' ability to engage in economic and political organisations (Arrow (1974), Hofstede, Hofstede, and Minkov (2010)). In particular, it supports the creation of growth opportunities in countries characterised by weak institutions and informal economic activities (Alesina and La Ferrara (2002), Nunn (2014)). Hence, the study of the determinants of social cohesion appears fundamental in understanding growth paths of emerging markets. This chapter adds evidence to the literature on social cohesion in developing countries by exploring its relationship with agricultural practices. Indeed, as Litina (2016) highlights, the development of nations has been in part driven by the trade-off between land productivity and the associated levels of cooperation and social capital.

Social cohesion is here defined here as the prevalence of two societal char-

acteristics: collective preferences and the ability to bargain in favour of collectivism (Alan et al. (2021), Mendez, Otero, Link, Morales, and Gayo (2021)). Recent studies on the origins of collectivism show historical trends that link changes in agricultural practices to changes in collectivism in different societies (Buggle (2017), Fiszbein et al. (2022)). In particular, the economic history literature hypothesises that the prevalence of labour-intensive agriculture fostered over the years cooperation due to the repeated interactions between stable and dedicated groups (Schulz, Bahrami-Rad, Beauchamp, and Henrich (2019), Fiszbein et al. (2022), Vollrath (2011)). Similarly, Talhelm, Zhang, Oishi, Shimin, Duan, Lan, and Kitayama (2014) show that even within the same country, i.e. China, a different history of farming across regions strongly affected the culture of their residents in terms of cooperative attitudes.

Compared to the existing literature that relies on proxies to measure collectivism (e.g. Hofstede et al. (2010), Buggle (2017), Fiszbein et al. (2022)), this work has the advantage of having a direct measure for collectivism and social cohesion, observing individuals' behaviour in public goods games in 70 Colombian municipalities (Attanasio et al. (2012)).

Merging the experimental data with CEDE Municipal Panel data, the empirical analysis shows that agricultural productivity and the share of hectares devoted to labour-intensive crops are positively correlated to the level of social cohesion registered in municipalities part of the experiment.

To explore the causality of this relationship, since the prevalence of labour-intensive crops might suffer from endogeneity issues related to individual preferences and technology availability (Fiszbein et al. (2022)), I use an instrumental variable methodology, instrumenting the share of labour-intensive crops with the average annual rainfall from 1989 to 2006 (the year the experiment was implemented), under the assumption that the historical levels of precipita-

tions directly affected soil moisture and crop suitability, influencing, together with temperature, altitude and soil composition, the type of crops harvested in each region (Fischer et al. (2002), Litina (2016), Buggle and Durante (2021)).

The IV results show that a one-standard-deviation rise in the average annual rainfall in a municipality increases the share of collective preferences by 27% and, by 33% the ability to bargain to the community's advantage. Results add interesting evidence to the links between labour market characteristics and social cohesion, showing that labor-intensive production systems can positively affect social cohesion and, as a consequence, development opportunities in emerging markets.

The rest of the paper is organised as follows. Section 3.2 details the data collected and the measures for agriculture and social cohesion. Section 3.3 describes the empirical strategy adopted to study the relationship between agricultural practices and social cohesion. Lastly, Section 3.4 reports the results of the empirical analysis, while Section 3.5 discusses the assumptions of the IV strategy implemented.

3.2 Data

The data used in the analysis are collected from three sources. First, measures of social cohesion are collected through a lab-in-the-field experiment implemented in 2006 in Colombia[1]. Second, data on municipalities characteristics are collected and distributed by the research center CEDE at Universidad de los Andes[2]. Lastly, data on historical precipitations are collected by the Euro-

[1]O. Attanasio, A. Barr, J.C. Cardenas, G. Genicot, C. Meghir, `https://www.openicpsr.org/openicpsr/project/113814/version/V1/view`, accessed Mar 20, 2022.

[2]CEDE Municipal Panel, Centro de Estudios sobre Desarrollo Economico (CEDE), Universidad de Los Andes, Colombia, `https://datoscede.uniandes.edu.co/es/catalogo-de-microdata`, accessed May 11, 2023.

pean Center for Medium-Range Weather Forecasts and distributed by FAO[3]. All the above mentioned data and measures created for the empirical analysis are described in detail in the current section.

3.2.1 Social cohesion measures

To measure social cohesion, individual data are collected through a lab-in-the-field experiment implemented in 2006 in Colombia by Attanasio et al. (2012). The collected information is used to measure the two variables of interest in studying social cohesion: the existence of collective preferences and the ability to bargain in favour of collectivism.

The laboratory experiment involves around 2,300 participants selected randomly in 70 Colombian municipalities. Municipalities are part of the evaluation of the *Familias en Acción* program (see Attanasio et al. (2012) for more details on the experimental design and Attanasio, Battistin, and Mesnard (2012) for information on the program). A balance table on the characteristics of the selected municipalities is reported in Table 3.5.

Participants in the experiment are asked to play two public goods (PG) game rounds. The PG game consists of 40 people given a token valued at 2,000 Colombian pesos and individually asked to allocate it simultaneously to a private or a public account. In the latter case, the value of the token doubles. After the private decision is made, each participant receives the final value of their token together with the shared value of the tokens allocated to the public account by other players.

By construction, the expected return from cooperation to the PG game increases with the number of expected people contributing (n^E). However, since revenues from the public account are also shared among those who decide

[3]Country Indicators, FAO - Agricultural Stress Index System (ASIS), `https://www.fao.org/giews/earthobservation/country/index.jsp?lang=en`, accessed May 25, 2024.

not to cooperate, the expected return from cooperation is always lower than the expected return from non-cooperation. Indeed, for any given n^E:

$$E[public_{account}] = \frac{4000}{40} + n^E \frac{4000}{40} < E[private_{account}] = 2000 + n^E \frac{4000}{40}$$

As Kurzban and Houser (2011) report, the existing literature reports a strong correlation between players' contributions and their expectations about others' contributions. Moreover, as shown by Charness et al. (2014), contributions to PG games are expected to be higher in settings where a group identity exists. For this reason, round 1 individual choice is used to measure the existence of collective preferences.

After playing round 1, participants can talk to each other for 15 minutes before playing the same game for the second, and last, round. Given the possibility of discussion with other players, round 2 behaviour, and its deviations from round 1, is used to study the bargaining process. In particular, a higher number of individuals contributing in round 2, controlling for round 1 contributions, is considered a good indicator of a social bargaining ability in favour of collectivism.

Table 3.1 shows the overall number of players by contribution choice in each round, while Figure 3.1 shows the share of players contributing in each round by municipality. The share of people contributing in round 2 slightly increases from round 1 in most municipalities (Figure 3.1). Indeed, as reported in Table 3.1, the total number of people contributing to round 2 is higher than the number of people contributing in round 1, with 392 people switching choices from "private" to "public" account across rounds and only 243 people switching from "public" to "private" account.

It is worth noticing that this might be correlated to the prevalence of women in the sample (Table 3.2). Indeed, as shown by Bram Cadsby and Maynes (1998), women are significantly better at coordinating around a selected equilibrium in PG games compared to men. Yet, overall 45% of players (948) never contributed to the PG game, relative to 25% of players (529) who chose to contribute in both rounds.

As Table 3.2 shows, participants in the experiment are primarily women and with low education. On average, 78% of them are married and live in rural Colombia. The number of family ties in each group ranges between 0 and 5, with an average number of 0.32 family members recognised, while the number of friendship ties ranges between 0 and 16, with an average number of friends recognised of 2.42.

As reported in Table 3.3, on average less than 10% of people are linked by family or friendship ties in the PG games, with a lot of variation across municipalities. The average share of acquaintances is slightly higher than the share of family or frienship ties, but still very low (15%). Moreover, on average less than half of the people contributed to both rounds of the PG game. The standard deviation of contributions to both rounds is around 0.5, with the median number of people contributing to the PG game (both first and second rounds) equal to 0.

Table 3.1: Number of players by contribution choice

	Round 2	
Round 1	Private	Public
Private	948	392
Public	243	529
Observations	2112	

Table 3.2: Characteristics of participants to the experiment

	mean	median	min	max	standard deviation
Demographic characteristics					
Female	0.87	1	0	1	0.33
Age	41.72	40	16	84	11.37
Education = years of schooling	3.7	3	0	17	3.13
HH heads	0.29	0	0	1	0.45
Married	0.78	1	0	1	0.42
Live in municipal centre	0.34	0	0	1	0.48
Network characteristics					
Number of family ties recognised	0.32	0	0	5	0.67
Number of friendship ties recognised	2.42	2	0	16	2.57
Observations	2321				

Table 3.3: Characteristics of PG game

	mean	median	min	max	standard deviation
Contribution 1st round	0.37	0	0	1	0.48
Contribution 2nd round	0.44	0	0	1	0.5
N. players	35.93	40	11	40	6.1
Share of family members	0.01	0	0	0.19	0.02
Share of friends	0.07	0.05	0	0.5	0.07
Share of acquaintances	0.15	0.1	0	0.92	0.17
Observations	2321				

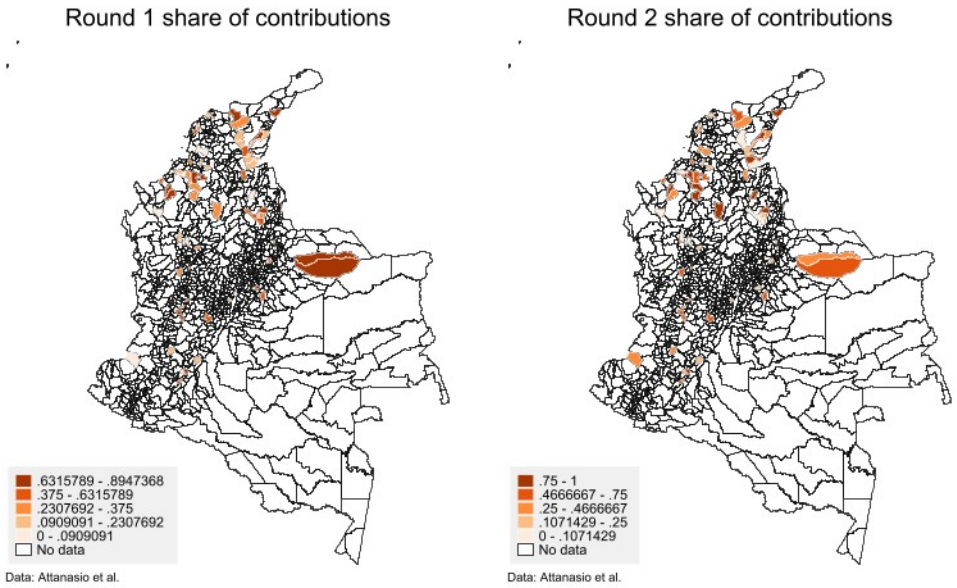


Figure 3.1: Average PG game contributions

3.2.2 Municipalities characteristics

Data on municipalities' characteristics are extracted from the CEDE Municipal Panel datasets. For each municipality part of the laboratory experiment, the following information is collected:

- Agricultural production = Tons of agricultural production by crop seeded (2007 data)
- Agricultural hectares = Hectares by crop seeded (2007 data)
- Altitude = altitude in meters for each municipality (2006 data)
- Distance from the major city = distance in km from the closest capital of the district (2006 data)
- Total acts of terrorism in 2005 = total number of terroristic attacks to the community (2005 data)

Later in this section, the differences in data collection years are explained. For the empirical analysis, the following two variables are calculated to measure agricultural productivity and the share of labour-intensive agriculture:

- Agricultural total productivity = Tons/Hectare of agricultural production (2007 data)
- Hectares share of labour-intensive crops = share of hectares of total agricultural production devoted to labour-intensive crops (2007 data)

The definition of labour-intensive crops comes from the World Bank report on agricultural and rural competitiveness (WorldBank (2003)), which identifies the five most highly labour-intensive crops in Colombia: isabella grape, green asparagus, coffee, granadilla and tobacco (Figure 3.2).

Due to data availability, information on agricultural activities refer to 2007, the year after the experiment implementation. Although productivity might change yearly, the share of hectares devoted to labour-intensive crops should not suffer from significant variations across years. Hence, the different years of data collection is not expected to affect the results of the analysis.

Figure 3.3 shows the geographical distribution of the two variables (agricultural productivity and hectares share of labour-intensive crops) in all Colombian municipalities, while Figure 3.4 shows the distribution in the 70 municipalities part of the experiment.

The variables altitude, distance and acts of terrorism are used as municipality-level controls in the analysis because they might confound the relationship between agricultural practices and social cohesion. Indeed, altitude and distance can affect social isolation and collective preferences. On the other hand, the history of Colombia has led to a strong relationship between terroristic acts and individuals' trust and, consequently, social capital (Rubio-Ramos (2024)).

For this reason, the variable *total acts of terrorism* refers to 2005 data, i.e. the year before the experiment is implemented, in order to capture possible variations in the frequency of terroristic attacts in the municipalities under study.

Table 3.4 summarises the main characteristics of the municipalities selected in the laboratory experiment. Municipalities are primarily rural areas. However, a lot of variation is registered in terms of location. Some municipalities are closer to the major city of each department (14 km), while some are located very far away (more than 280 km). Also, the distribution in terms of altitude varies a lot, ranging from 3 to 3,000 metres above the sea level. Municipalities were randomly selected. Hence, as shown in Table 3.5, municipalities part of the sample exhibit characteristics that are not statistically different than the rest of the country.

In terms of agricultural productivity, the least productive municipality reports 0 tons per hectare produced in 2007, while the most productive reports 645 tons per hectare of agricultural production. Instead, regarding hectares devoted to labour-intensive crop production, municipalityies' share spans from 0 to 85%. Figure 3.3 shows the distribution of agricultural productivity and hectares share of labour-intensive crops on the whole Colombian land, while Figure 3.4 reports the same distributions in the municipalities part of the experiment.

Table 3.4: Characteristics of selected municipalities

	mean	median	min	max	sd
Altitude (m)	724.53	249	3	3269	898.77
Distance from the major city	75.75	59.98	13.79	286.17	57.56
Total acts of terrorism in 2005	0.27	0	0	6	0.94
Ton/Hectares - agriculture	105.28	86.29	10.36	313.95	75.22
Hectare share of labour-intensive crops	0.16	0	0	0.85	0.25
Observations	68				

Crop	Labor Use	Crop	Labor Use	Crop	Labor Use
Isabella grapes	520.0	*Lulo*	116.4	White corn	39.8
Green asparagus	339.4	Potatoes	104.6	Oil palm	35.2
Coffee	249.7	Plantain	97.4	Irrigated rice	33.4
Granadilla	220.8	Citrus	96.9	Yellow corn	23.3
Light tobacco	214.2	Industrial cassava	87.9	Rainfed rice	22.0
Pitahaya	208.9	Mango	84.4	Wheat	20.6
Dark tobacco	196.3	Beans	75.9	Soybeans	18.0
Passion fruit	166.2	Sugarcane	65.1	Barley	15.0
Cacao	121.7	Cotton	49.9	Sorghum	13.4
Papaya	121.1				

Source: WorldBank (2003).

Figure 3.2: Labor use by crop in Colombia

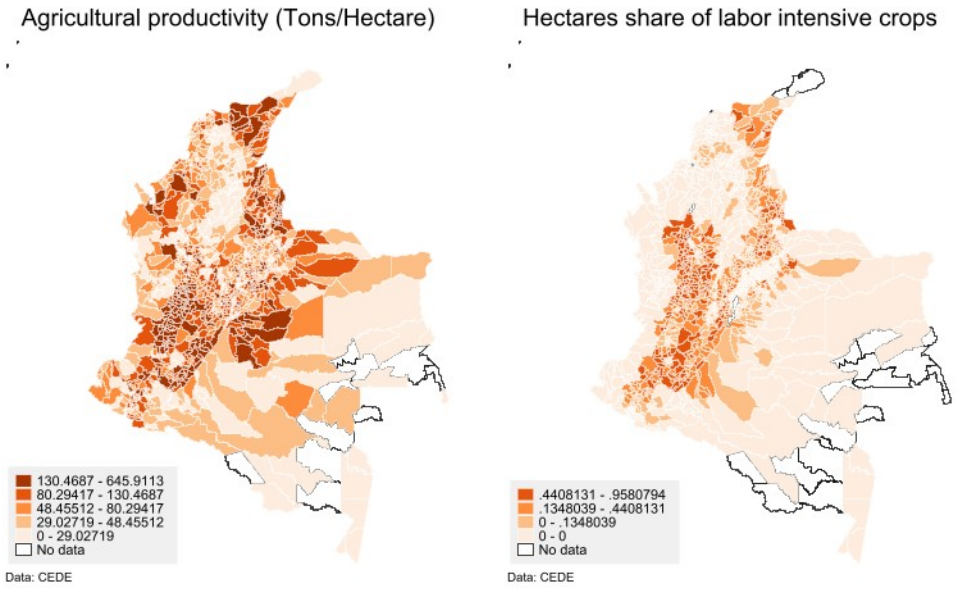


Figure 3.3: Municipalities data - all municipalities

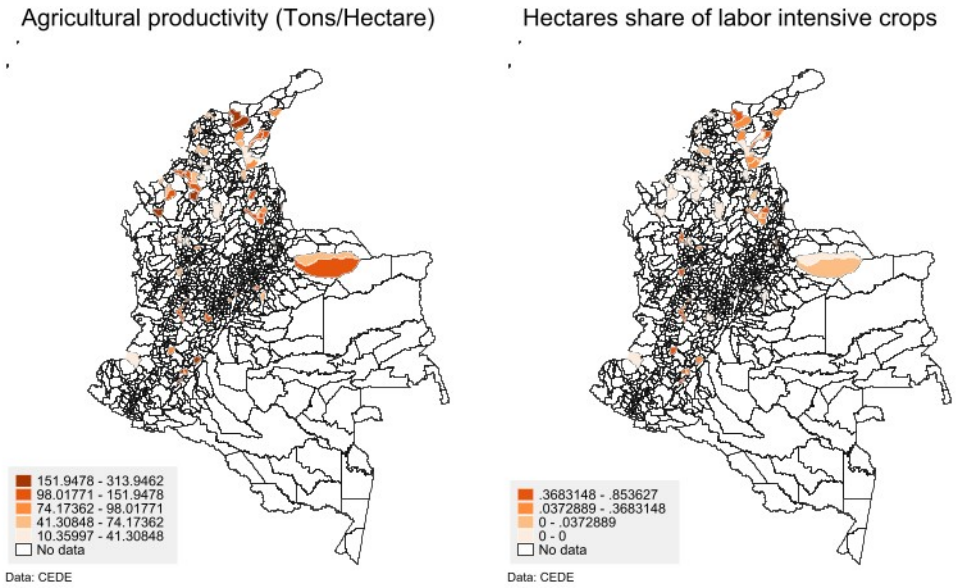


Figure 3.4: Municipalities data - selected municipalities

3.2.3 Rainfall data

To implement the IV strategy, the last piece of information collected is historical precipitation data, forecasted by the European Center for Medium-Range Weather Forecasts. Using the rainfall volumes (in mm) for each dekad from 1989-2024, the following variable is calculated: the average yearly cumulative rainfall volumes (in meters) for the almost two decades before the experiment's implementation (1989-2006).

As Fischer, Nachtergaele, Prieler, Teixeira, Teth, van Velthuizen, Verelst, and Wiberg (2012) highlight, the relationship between annual rainfall and agriculture is not linear. Indeed, to guarantee adequate soil moisture, precipitations should ensure a minimum level of water availability, but excess rain can cause floods and damage crops. Hence, to capture the non-linearity of the relationship, the following instrumental variables are used in the empirical analysis:

- $Rainfall$ = average yearly cumulative rainfall volumes (1989-2006)

- $Rainfall^2$ = quadratic value of the variable $Rainfall$

Figure 3.5 shows the distribution of annual precipitations. The forecasted rainfall volumes in the municipalities part of the lab experiment range from 1,400mm to around 5,600mm of yearly rainfall, while the national distribution of average precipitations ranges from 1,400mm to almost 9,000mm of annual rainfall. Hence, in the analysis there are no municipalities in the upper extreme of the overall national distribution, i.e. those that experience excessive rain.

On average, selected municipalities exhibit the same charateristics compared to excluded ones. Indeed, as shown in Table 3.5 the variables selected for the analysis have mean values that do not differ significantly across samples.

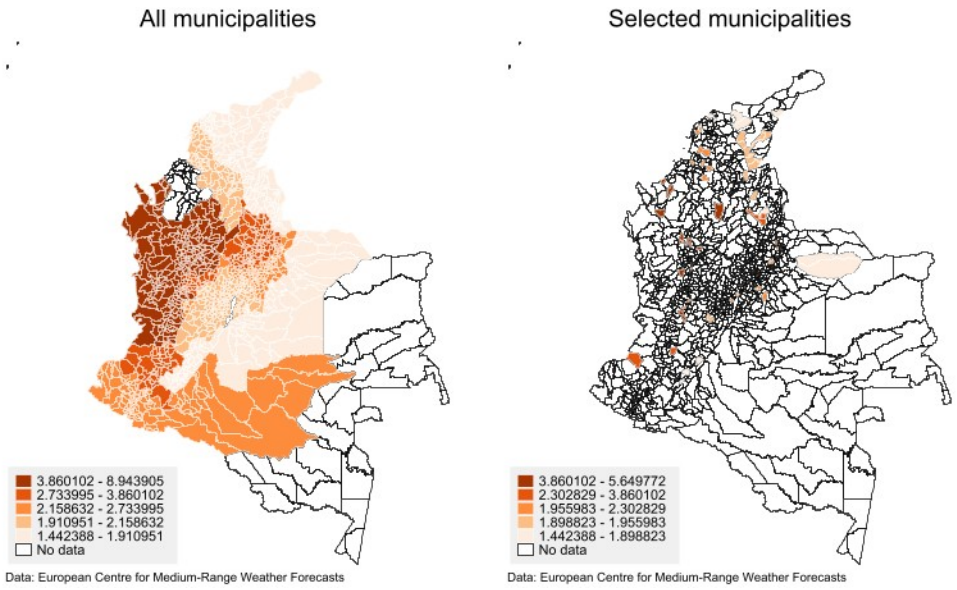


Figure 3.5: Average yearly precipitations 1989-2006

Variable	(1) Excluded N	(1) Excluded Mean/(SE)	(2) Included N	(2) Included Mean/(SE)	(1)-(2) Pairwise t-test N	(1)-(2) Pairwise t-test Mean difference
Distance from the major city	1054	81.826	68	75.746	1122	6.080
		(1.872)		(6.980)		
Total acts of terrorism in 2005	1054	0.564	68	0.265	1122	0.299
		(0.063)		(0.114)		
Ton/Hectares - agriculture	1042	88.479	68	105.281	1110	-16.801
		(2.690)		(9.122)		
Hectare share of labour-intensive crops	1028	0.192	68	0.155	1096	0.037
		(0.008)		(0.030)		
Average yearly precipitations 1989-2006	998	3.047	61	2.726	1059	0.322
		(0.051)		(0.166)		
F-test of joint significance (F-stat)						1.093
Observations						1055

Table 3.5: Balance table on municipalities characteristics

3.3 Empirical strategy

The empirical strategy aims to isolate the impact of labour-intensive agriculture on social cohesion. The estimated model is the following:

$$\text{social cohesion}_i = \alpha + \beta \text{agriculture}_m + \delta X_{i,m} + \epsilon \tag{3.1}$$

where i = player i and m = municipality m.

The dependent variable *social cohesion* is estimated using the results of both rounds of the PG games to capture the characteristics of social cohesion we are interested in: the existence of collective preferences and the ability to bargain in favour of collectivism (see Section 3.2.1 for the details on the variables used).

Collective preferences are measured using the outcome results from round 1:

$$\textit{contribution round 1} = \begin{cases} 1 & \text{if the player contributed to the public good} \\ 0 & \text{if the player kept the token private} \end{cases}$$

Indeed, as anticipated in Section 3.2.1, when choosing to contribute in round 1, individuals' contribution choice is based on the expectations about other people from the same municipality part of the experiment (Kurzban and Houser (2011)). Hence, round 1 contributions capture the overall expectations about the group's willingness to cooperate and the presence of collective preferences.

Instead, in round 2, the decision to contribute depends on adjustments of beliefs from round 1 results and the bargaining between rounds. To capture the second factor, i.e., the ability to bargain in favour of collectivism, round 2 contributions are analysed, controlling for round 1 behaviour and beliefs

updates about other players' behaviour (i.e. share of players contributing in round 1):

- *contribution round 2* = $\begin{cases} 1 & \text{if the player contributed to the public good} \\ 0 & \text{if the player kept the token private} \end{cases}$
- *% contributions round 1* = $\frac{\text{n. players contributing in round 1}}{\text{n. players}}$

On the other hand, the independent variable *agriculture* is measured using CEDE data on agricultural productivity and hectares share of labour-intensive crops (see Section 3.2.2 for more details on the measures):

- *Log Ton/Hectares* = $\log(\frac{\text{tons agricultural production}}{\text{hectares agricultural production}})$
- *% Hectares - 5 crops* = $\frac{\text{hectares 5 most labor intensive crops}}{\text{hectares agricultural production}}$

The first step of the analysis explores the correlation between the two variables of interest: agriculture and social cohesion. A probit regression is implemented on the individual choices during the PG games[4], using as controls the gender, education and marital status of each player, the number of family and friends recognised during the game, acts of terrorism registered in the municipality in the previous year, distance of the municipality from the major city, and municipality altitude (see Section 3.2 for more details on the confounding factors). Estimated results are shown in Section 3.4.

As Buggle (2017) and Fiszbein et al. (2022) note, OLS results might be upward biased due to endogeneity in the type of crops harvested for two reasons. First, reverse causality may emerge if historical migration has attracted individuals with an attitude towards collectivism versus individualism that has shaped the type of labour activities performed. The same is true regarding the

[4]Two municipalities are dropped from the analysis because of the lack of variation in contributions to the PG game.

availability of technology that might have affected the kind of crops cultivated. Second, there can be an omitted variables bias. For these reasons, the second step of the analysis relies on an instrumental variable (IV) methodology to exclude the possible endogeneity from the study.

The variable used in the IV strategy is the average yearly rainfall from 1989-2006:

$$Rainfall = \frac{\text{cumulative rainfall volume (m)}}{\text{n. years}}$$

The intuition is that historical precipitations affect the type of crops produced through soil moisture and water availability (Fischer et al. (2012), Amare, Jensen, Shiferaw, and Cisse (2018), Ceballos-Sierra and Dall'Erba (2021)). The exclusion restriction relies on the argument that, conditional on the geographic covariates included, average historical precipitations are plausibly exogenous to the level of social cohesion in each municipality (Rubio-Ramos (2024)). Further discussion on the assumptions is reported in Section 3.5.

As anticipated in Section 3.2.3, since Fischer et al. (2012) show that the relationship between annual rainfall and agriculture is not linear, the regression includes average yearly rainfall as well as its squared value to instrument labour-intensive agriculture. Intuitively, precipitations should ensure a minimum level of water availability to guarantee adequate soil moisture. However, excess rain can cause floods and still damage the crops (see Section 3.2.3 for more details).

The IV strategy focuses on the effects of labour intensive agriculture (*% Hectares - 5 crops*) on social cohesion and relies on a 2SLS model. The first stage is identified as:

$$\widehat{agriculture}_m = \lambda_1 + \lambda_2 rainfall_m + \lambda_2 rainfall_m^2 + \lambda_4 X_{i,m} + \phi \qquad (3.2)$$

The second stage is identified as:

$$\text{social cohesion}_i = \alpha + \beta \widehat{agriculture}_m + \delta X_{i,m} + \epsilon \tag{3.3}$$

where, i = player i and m = municipality m.

Results of the estimated equations are reported in the next section. Coefficients for equation (3.1) are reported in Table 3.6 using as independent variable *Log Ton/Hectares - agriculture* and in Table 3.7 using as independent variable *% Hectares - 5 crops*. Table 3.8, instead, reports the results from equations (3.2) and (3.3) using the IV strategy.

3.4 Results

Table 3.6 reports the regression results from equation (3.1) using the log value of tons per hectare of agricultural production to measure the variable *agriculture*, which includes all crops produced in each municipality (variable *Log Ton/Hectares*).

Results show that both variables of *social cohesion* are positively and significantly correlated with the variable *agriculture*. In particular, a one standard deviation of log agricultural productivity is associated with a 0.11 points increase in round 1 contribution and with a 0.21 points increase to round 2 contribution, in the latter case controlling for both individual collective contributions to round 1.

The same relationship is reported in Table 3.7, using as a measure for the variable *agriculture* the share of hectares devoted to the five most labour-intensive crops produced in Colombia[5] on the total number of hectares devoted to agriculture (variable *% Hectares - 5 crops*). The coefficients confirm the

[5]WorldBank (2003).

positive correlation between agriculture and social cohesion, even bigger in magnitude compared to Table 3.6 coefficients: a coefficient of 0.58 for contributions to round 1 (compared to 0.11 in Table 3.6) and a coefficient of 0.42 for contributions to round 2 (compared to 0.21 in Table 3.6).

Lastly, the IV regression results (Table 3.8) confirm the upward bias of the OLS coefficients and show evidence about the causality of the relationship between agricultural prevalence and social cohesion. Indeed, the coefficient capturing the relationship between agriculture and first round contributions is positive and significant at a 10% level, reporting a higher likelihood of collective preferences in municipalities with a higher share of labour-intensive agriculture (Table 3.8, Column (1)). Instead, the instrumented coefficient for round 2 contributions, controlling for individual and average contributions in round 1, is positive and statistically significant at a 1% level and reports an increase of 33% in the ability to bargain for collectivism (Table 3.8, Column (2)).

Overall results confirm the hypothesis that a labour-intensive production system has, over time, shaped the societal structure in each municipality by boosting informal cooperation and social cohesion. Indeed, the results appear stronger when using the five most labour-intensive crops, compared to agricultural productivity, as a measure for the variable *agriculture*.

In particular, when instrumenting the variable *agriculture*, labour-intensive production appears more influencial on the societal ability to bargain in favour of cooperation (Table 3.8, Column (2)), with a still significant but lower effect on the existence of collective preferences (Table 3.8, Column (1)). The coefficients of the other covariates included in the analysis are reported in the Appendix 3.7.

VARIABLES	(1) Contribution round 1	(2) Contribution round 2
Log Ton/Hectares	0.109***	0.212***
	(0.0414)	(0.0453)
Contribution round 1		0.770***
		(0.0730)
% Contributions round 1		0.778***
		(0.146)
Constant	-2.044***	-3.801***
	(0.309)	(0.373)
Controls	Yes	Yes
Observations	2,112	2,112

Table 3.6: Correlation between social cohesion and agricultural productivity

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The table shows equation (3.1) coefficients results using a probit regression and *Log ton/hectares* as the measure for *agriculture*. The variable *contribution round 1* captures the individual contribution choice in round 1, while the variable *% contributions round 1* indicates the share of players contributing in round 1 in the municipality. Column (1) reports the coefficients results on the dependent variable *contribution round 1*, which measures the existence of collective preferences, while column (2) reports the coefficients results on the dependent variable *contribution round 2*, which measures the bargaining ability. The controls included gender, marriage status, education, family and friends ties recognised by the player, number of players in the game, total acts of terrorism in 2005, distance from major cities and municipality altitude. Coefficients on the other covariates are reported in the Appendix 3.7.

VARIABLES	(1) Contribution round 1	(2) Contribution round 2
% Hectares - 5 crops	0.582***	0.422***
	(0.131)	(0.148)
Contribution round 1		0.766***
		(0.0729)
% Contributions round 1		0.778***
		(0.147)
Constant	-1.663***	-2.826***
	(0.231)	(0.259)
Controls	Yes	Yes
Observations	2,112	2,112

Table 3.7: Correlation between social cohesion and labor-intensive agriculture

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The table shows equation (3.1) coefficients results using a probit regression and *% hectares - 5 crops* as the measure for *agriculture*. The variable *contribution round 1* captures the individual contribution choice in round 1, while the variable *% contributions round 1* indicates the share of players contributing in round 1 in the municipality. Column (1) reports the coefficients results on the dependent variable *contribution round 1*, which measures the existence of collective preferences, while column (2) reports the coefficients results on the dependent variable *contribution round 2*, which measures the bargaining ability. The controls included gender, marriage status, education, family and friends ties recognised by the player, number of players in the game, total acts of terrorism in 2005, distance from major cities and municipality altitude. Coefficients on the other covariates are reported in the Appendix 3.7.

VARIABLES	(1) *Second stage* Contribution round 1	(2) *Second stage* Contribution round 2
% Hectares - 5 crops	0.271*	0.333***
	(0.141)	(0.129)
Contribution round 1		0.288***
		(0.024)
% Contributions round 1		0.329***
		(0.052)
Constant	-0.116	-0.296***
	(0.087)	(0.078)
Controls	yes	yes
Observations	1,844	1,844
R-squared	0.060	0.282
F test Cragg-Donald	120.2	116

Table 3.8: The effect of labor-intensive agriculture on collective preferences

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The table shows the coefficients of equation (3.3) using an instrumental variable regression. In column (1), the dependent variable *social cohesion* is measured with contributions to round 1 of the PG game (*contribution round 1*), while in column (2) it is measured with contributions to round 2 of the PG game (*contribution round 2*). The variable instrumented is *% hectares - 5 crops*, and the instruments are $rainfall$ and $rainfall^2$. In column (2), the variable *% contributions round 1* indicates the share of players contributing in round 1 in the municipality. The controls include gender, marriage status, education, family and friends ties recognised by the player, number of players in the game, total acts of terrorism in 2005, distance from major cities and municipality altitude. Coefficients on the other covariates are reported in the Appendix 3.7. See Table 3.9 for the first stage results.

3.5 Instrument assessment

In this section, an assessment of the instrument is performed, discussing the assumption underlying the IV design, which exploits the relationship between historical precipitation data and the type of crops cultivated in Colombian municipalities. In particular, a good instrument should be relevant and valid. The following two subsection will expand on these two, while the last subsection explores how the monotonicity assumption applies to this setting.

3.5.1 Instrument relevance

The first stages of the IV regressions show, in accordance to predictions, that an increase precipitation levels correlates significantly and non-monotonically with the share of labor intensive crops cultivated in the municipality (Table 3.9). In particular, the coefficient of the variable $Rainfall$ shows that, controlling for its squared value ($Rainfall^2$), an increase in precipitation levels correlates with a decrease in the share of labor intensive crops cultivated. Overall, however, the relationship is convex, and follows a U-shaped functional form. Results hold for both variables of *social cohesion* (i.e. round 1 contribution reported in Column (1), and round 2 contribution reported in Column (2)). F-statistics are above the conventional threshold for weak instruments.

3.5.2 Instrument validity

As anticipated in Section 3.3, the exclusion restriction is based on the argument that average historical precipitations are plausibly exogenous to the level of social cohesion (Rubio-Ramos (2024)). The condition would be violated if precipitations capture factors that affect social cohesion through channels different than agriculture. This is a plausible hypothesis when we deal with

precipitation shocks, which can have an effect on social cohesion by boosting reciprocal support in periods of distress. However, the use of historical average precipitations aims at smoothing these effects, by isolating the expected amount of yearly precipitations. Indeed, as reported by Fischer et al. (2012), Amare et al. (2018) and Ceballos-Sierra and Dall'Erba (2021), cultivation choices are strongly dependent on expectations about yearly and seasonal precipitations, that in turn affect soil moisture and crops suitability.

3.5.3 Monotonicity

In this setting, the monotonicity assumption requires that municipalities that are less suited for labor intensive crops cultivations, would also be more social cohese if they were more suited to growing labor intensive crops. To provide evidence in favour of this assumption, Table 3.10 shows that for extreme quartiles of average precipitation levels, the first stage estimates are statistically significant and consistent with the coefficients signs reported in Table 3.9.

	(1)	(2)
	First stage	*First stage*
VARIABLES	% Hectares - 5 crops	% Hectares - 5 crops
Rainfall	-0.277***	-0.206***
	(0.035)	(0.037)
Rainfall^2	0.050***	0.040***
	(0.005)	(0.006)
F test Cragg-Donald	120.2	116
R-squared	0.241	0.255
Observations	1,844	1,844

Table 3.9: Correlation between precipitations and agriculture

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The table shows the results from equation (3.2). Column (1) uses as the second stage dependent variable round 1 contributions to the PG game, while column (2) uses round 2 contributions to the PG game. Controls include gender, marriage status, education, family and friends ties recognised by the player, number of players in the game, total acts of terrorism in 2005, distance from major cities and municipality altitude. In column (2) controls on individual and collective round 1 contributions are added. See Table 3.8 for the second stage results.

Dependent variable:	(1)	(2)
% Hectares - 5 crops		
Sample	Rainfall	Rainfall
	(25^{th} percentile)	(75^{th} percentile)
Rainfall	-18.90***	-1.544***
	(2.948)	(0.509)
Rainfall^2	5.713***	0.169***
	(0.868)	(0.0559)
Constant	15.43***	4.004***
	(2.456)	(1.161)
Observations	497	384
R-squared	0.606	0.498

Table 3.10: Monotonicity test

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The table shows the first stage of the IV approach for different population subsamples.

3.6 Conclusions

This chapter analyses the relationship between labour-intensive agriculture and social cohesion in developing countries. Using a detailed and innovative dataset on cooperative attitudes in Colombian municipalities, this work shows that agricultural practices can shape collective preferences and the ability to cooperate among individuals in a society.

The empirical strategy relies on a lab-in-the-field experiment implemented in 70 municipalities in Colombia, which collects information on the individual likelihood of contributing to public goods games. Observed behaviours are then used to measure social cohesion, defined as the presence of collective preferences and the ability to bargain in favour of collectivism.

To avoid endogeneity issues in the empirical analysis, an IV strategy is implemented, instrumenting the share of hectares devoted to labour intensive crops with historical data on precipitations. The intuition behind the instrument relies on the direct relationship between rainfall and soil suitability for crop production (Fischer et al. (2012)).

Results show a positive and significant effect of labour intensive agriculture on social cohesion, both on the presence of collective preferences and on the bargaining ability. This adds relevant evidence to the literature on the relationship between historical agricultural practices and current preferences (Buggle (2017), Fiszbein et al. (2022), Litina (2016), Talhelm et al. (2014)) by providing a detailed analysis of individual behaviour in a laboratory setting, hence isolating the cohesive attitudes from confounding factors that might arise in the natural setting.

Moreover, it adds evidence to the literature on social cohesion (e.g. Alan et al. (2021), Mendez et al. (2021)) by highlighting how labour market characteristics can significantly affect the levels of prosocial behaviour at the mu-

nicipality level, notwithstanding the existing network ties.

Existing historical evidence on the links between agricultural practices and collectivism in different societies (Fiszbein et al. (2022), Vollrath (2011), Rodriguez-Giraldez and Vicente-Ramos (2020)) suggest that the results of the analysis have a high external validity and can help better understanding the determinants of social cohesion.

3.7 Appendix

Regressions controls coefficients

	(1)	(2)
VARIABLES	Contribution PG round 1	Contribution PG round 2
Female	-0.063	0.017
	(0.087)	(0.089)
Education	0.009	-0.006
	(0.009)	(0.01)
Married	0.069	-0.011
	(0.069)	(0.075)
Family ties in the game	0.064	0.025
	(0.042)	(0.044)
Frienship ties in the game	0.033***	0.007
	(0.012)	(0.012)
Number of players in the game	0.029***	0.052***
	(0.005)	(0.006)
Total acts of terrorism in 2005	-0.217***	-0.019
	(0.039)	(0.041)
Distance from the major city	0.002***	0.005***
	(0.001)	(0.001)
Altitude	-6e-05	-0.0002***
	(4e-05)	(4e-05)
Observations	2,112	2,112

Table 3.11: Correlation between social cohesion and agricultural productivity: controls coefficients

VARIABLES	(1) Contribution PG round 1	(2) Contribution PG round 2
Female	-0.047	0.046
	(0.086)	(0.089)
Education	0.01	-0.004
	(0.009)	(0.01)
Married	0.07	-0.007
	(0.069)	(0.074)
Family ties in the game	0.051	0.002
	(0.042)	(0.044)
Frienship ties in the game	0.038***	0.008
	(0.012)	(0.012)
Number of players in the game	0.03***	0.05***
	(0.005)	(0.006)
Total acts of terrorism in 2005	-0.227***	-0.046
	(0.039)	(0.041)
Distance from the major city	0.002***	0.005***
	(0.001)	(0.001)
Altitude	-0.0001***	-0.0003***
	(4e-05)	(5e-05)
Observations	2,112	2,112

Table 3.12: Correlation between social cohesion and labor-intensive agriculture: controls coefficients

VARIABLES	(1) *Second stage* Contribution round 1	(2) *Second stage* Contribution round 2
Female	-0.009	-0.020
	(0.033)	(0.029)
Education	0.005	-0.001
	(0.004)	(0.003)
Married	0.025	-0.010
	(0.026)	(0.023)
Family ties in the game	0.017	0.014
	(0.016)	(0.014)
Frienship ties in the game	0.011**	0.016***
	(0.005)	(0.004)
Number of players in the game	0.011***	0.009***
	(0.002)	(0.002)
Total acts of terrorism in 2005	-0.058***	0.009
	(0.012)	(0.011)
Distance from the major city	0.001***	0.002***
	(0.0002)	(0.0002)
Altitude	-5e-05***	-6e-05***
	(2e-05)	(2e-05)
Observations	1,844	1,844

Table 3.13: The effect of labor-intensive agriculture on collective preferences: controls coefficients

Bibliography

Adler, N., G. Castellazzo, E. Epel, and J. Ickovics (2000). Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy, white women. *Health Psychology 19*(6), 586–592.

Akerlof, G. A. and R. E. Kranton (2000). Economics and identity. *The Quarterly Journal of Economics 115*(3), 715–753.

Alan, S., C. Baysan, M. Gumren, and E. Kubilay (2021). Building Social Cohesion in Ethnically Mixed Schools: An Intervention on Perspective Taking*. *The Quarterly Journal of Economics 136*(4), 2147–2194.

Alesina, A. and E. La Ferrara (2002). Who trusts others? *Journal of Public Economics 85*(2), 207–234.

Alesina, A., S. Stantcheva, and E. Teso (2018). Intergenerational mobility and preferences for redistribution. *American Economic Review 108*(2), 521–554.

Altonji, J. G., F. Hayashi, and L. J. Kotlikoff (1992). Is the extended family altruistically linked? direct tests using micro data. *The American Economic Review 82*(5), 1177–1198.

Amare, M., N. D. Jensen, B. Shiferaw, and J. D. Cisse (2018). Rainfall shocks and agricultural productivity: Implication for rural household consumption. *Agricultural Systems 166*, 79–89.

Arrow, K. (1974). *The Limits of Organization*. Fels Center of Government Series. Norton.

Attanasio, O., A. Barr, J. C. Cardenas, G. Genicot, and C. Meghir (2012). Risk pooling, risk preferences, and social networks. *American Economic Journal: Applied Economics 4*(2), 134–167.

Attanasio, O., E. Battistin, and A. Mesnard (2012). Food and cash transfers: Evidence from colombia. *Economic Journal 122*(559), 92–124.

Batson, C. (2010). Empathy-induced altruistic motivation. *Prosocial Motives, Emotions, and Behavior: The Better Angels of Our Nature*.

Batson, C. D. (1990). How social an animal? the human capacity for caring. *American Psychologist 45*, 336–346.

Benabou, R. and J. Tirole (2006). Belief in a just world and redistributive politics. *The Quarterly journal of economics 121*(2), 699–746.

Berg, J., J. Dickhaut, and K. McCabe (1995). Trust, reciprocity, and social history. *Games and Economic Behavior 10*(1), 122–142.

Bram Cadsby, C. and E. Maynes (1998). Gender and free riding in a threshold public goods game: Experimental evidence. *Journal of Economic Behavior & Organization 34*(4), 603–620.

Bramoulle, Y. and R. Kranton (2007). Risk-sharing networks. *Journal Of Economic Behavior and Organization 64*(3-4), 275–294.

Buggle, J. C. (2017). Irrigation, Collectivism and Long-Run Technological Divergence. Cahiers de Recherches Economiques du Departement d'economie 17.06, Universite de Lausanne, Faculte des HEC, Departement d'economie.

Buggle, J. C. and R. Durante (2021). Climate Risk, Cooperation and the Co-Evolution of Culture and Institutions. *The Economic Journal 131*(637), 1947–1987.

Ceballos-Sierra, F. and S. Dall'Erba (2021). The effect of climate variability on colombian coffee productivity: A dynamic panel model approach. *Agricultural Systems 190*, 103126.

Chan, J., H. To, and E. Chan (2006). Reconsidering social cohesion: Developing a definition and analytical framework for empirical research. *Social Indicators Research 75*(2), 273–302.

Charness, G., R. Cobo-Reyes, and N. Jimenez (2014). Identities, selection, and contributions in a public-goods game. *Games and Economic Behavior 87*, 322–338.

Dasgupta, D. and A. Saha (2022). Perceptions, biases, and inequality. *Journal of Economic Behavior & Organization 202*, 198–210.

Dearmon, J. and K. Grier (2009). Trust and development. *Journal of Economic Behavior & Organization 71*(2), 210–220.

Demuynck, T., P. Herings, R. Saulle, and C. Seel (2019). The myopic stable set for social environments. *Econometrica 87*(1), 111–138.

Evans, A. M. and J. I. Krueger (2009). The psychology (and economics) of trust. *Social and Personality Psychology Compass 3*(6), 1003–1017.

Exley, C. L. and K. Nielsen (2024). The gender gap in confidence: Expected but not accounted for. *American Economic Review 114*(3), 851–885.

Fafchamps, M. (2006). Development and social capital. *The Journal of Development Studies 42*(7), 1180–1198.

Fafchamps, M. and F. Gubert (2007). The formation of risk sharing networks. *Journal Of Development Economics 83*(2), 326–350.

Filippin, A. and M. Paccagnella (2012). Family background, self-confidence and economic outcomes. *Economics of Education Review 31*(5), 824–834.

Fischer, G., F. O. Nachtergaele, S. Prieler, E. Teixeira, G. Teth, H. van Velthuizen, L. Verelst, and D. Wiberg (2012). Global agro-ecological zones (gaez v3.0). Technical report, IIASA/FAO.

Fischer, G., H. van Velthuizen, M. M. Shah, and F. O. Nachtergaele (2002). Global agro-ecological assessment for agriculture in the 21st century. Technical report, IIASA/FAO.

Fiszbein, M., Y. Jung, and D. Vollrath (2022). Agrarian Origins of Individualism and Collectivism. NBER Working Papers 29643.

Foster, A. D. and M. R. Rosenzweig (2001). Imperfect commitment, altruism, and the family: Evidence from transfer behavior in low-income rural areas. *The Review of Economics and Statistics 83*(3), 389–407.

Glaeser, E. L., D. I. Laibson, J. A. Scheinkman, and C. L. Soutter (2000). Measuring Trust. *The Quarterly Journal of Economics 115*(3), 811–846.

Graafland, J. and B. Lous (2019). Income inequality, life satisfaction inequality and trust: A cross country panel analysis. *Journal of Happiness Studies*, 1717–1737.

Hofstede, G., G. Hofstede, and M. Minkov (2010). *Cultures and Organizations: Software of the Mind*. McGraw Hill LLC.

Jackson, M., T. Rodriguez-Barraquer, and X. Tan (2012). Social capital and social quilts: Network patterns of favor exchange. *American Economic Review 102*(5), 1857–1897.

Jakiela, P. and O. Ozier (2015). Does africa need a rotten kin theorem? experimental evidence from village economies. *The Review Of Economic Studies 83*(1), 231–268.

Kurzban, R. and D. Houser (2011). Individual differences in cooperation in a circular public goods game. *European Journal of Personality 15*, 37–52.

Lazzarini, S. G., R. C. Madalozzo, R. Artes, and J. O. Siqueira (2004). Measuring trust: An experiment in brazil. Insper working papers, Insper Working Paper, Insper Instituto de Ensino e Pesquisa.

Litina, A. (2016). Natural land productivity, cooperation and comparative development. *Journal of Economic Growth 21*(4), 351–408.

Lount, R. B. and N. C. Pettit (2012). The social context of trust: The role of status. *Organizational Behavior and Human Decision Processes 117*(1), 15–23.

Marmot, M. G. (2006). Status Syndrome: A Challenge to Medicine. *JAMA 295*(11), 1304–1307.

Mendez, M. L., G. Otero, F. Link, E. L. Morales, and M. Gayo (2021). Neighbourhood cohesion as a form of privilege. *Urban Studies 58*(8), 1691–1711.

Nunn, N. (2014). Chapter 7 - historical development. In P. Aghion and S. N. Durlauf (Eds.), *Handbook of Economic Growth*, Volume 2 of *Handbook of Economic Growth*, pp. 347–402. Elsevier.

Robinson, J. (2012). Limited insurance within the household: Evidence from a field experiment in kenya. *American Economic Journal: Applied Economics 4*(4), 140–164.

Rodriguez-Giraldez, W. and W. Vicente-Ramos (2020). Cultural aspects that influence the associative work of agricultural production chains in the mantaro valley of peru. *Management Science Letters 10*, 2425–2430.

Rubio-Ramos, M. (2024). Trust, violence, and coca. *Journal of Development Economics 167*, 103216.

Sapienza, P., A. Toldra-Simats, and L. Zingales (2013). Understanding trust. *The Economic Journal 123*(573), 1313–1332.

Schulz, J. F., D. Bahrami-Rad, J. P. Beauchamp, and J. Henrich (2019). The church, intensive kinship, and global psychological variation. *Science 366*(6466), eaau5141.

Siegrist, M. (2021). Trust and risk perception: A critical review of the literature. *Risk Analysis 41*(3), 480–490.

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American 223*(5), 96–103.

Talhelm, T., X. Zhang, S. Oishi, C. Shimin, D. Duan, X. Lan, and S. Kitayama (2014). Large-scale psychological differences within china explained by rice versus wheat agriculture. *Science 344*(6184), 603–608.

Townsend, R. M. (1995). Consumption insurance: An evaluation of risk-bearing systems in low-income economies. *The Journal of Economic Perspectives 9*(3), 83–102.

Vollrath, D. (2011). The agricultural basis of comparative development. *Journal of Economic Growth 16*(4), 343–370.

WorldBank (2003). Colombia: Agricultural and Rural Competitiveness. World Bank Publications - Reports 15711, The World Bank Group.