Universitat Autònoma de Barcelona

Essays on Development, Social Networks, and Information

Davide Pietrobo

Universitat Autònoma de Barcelona

Doctoral Thesis

Essays on Development, Social Networks, and Information

Author: Davide Pietrobon

Adviser: Tomás Rodríguez-Barraquer

Adviser: Pau Milán

A dissertation submitted to the *Departament d'Economia i d'Història Econòmica* in fulfillment of the requirements for the degree of Doctor of Philosophy by the International Doctorate in Economic Analysis (IDEA).

September 13, 2019

Acknowledgements

I thank the Fundació Markets, Organizations And Votes In Economics and Prof. Nezih Guner for the financial support provided to me during the Academic Year 2015-2016. I am grateful to the Ministerio de Ciencia y Innovación for the FPI scholarship BES-2015-075756. Thanks to these financial aids I was able to conduct the research presented in this thesis.

I am forever indebted to my advisers, Tomás Rodríguez-Barraquer and Pau Milán, for their guidance, mentorship, and friendship. I cannot count the hours they spent with me in their offices and cafes in Barcelona, passionately discussing every bit of every thought I had. Their presence has been invaluable not only for the quality of the research presented here but also for the quality of my life as a Ph.D. student. In my darkest hours, when my mind was wandering in space, they always managed to put me back on track and give me hope. I believe that this thesis would not have been possible without them.

I thank Fernando Vega-Redondo and Paolo Pin for hosting me at the Department of Economics at Bocconi University during the second semester of the Academic Year 2016-2017, and for providing useful insights during many conversations. I also thank Nicola Pavoni for his helpful suggestions.

I thank the faculty at IDEA, whose guidance was essential to this dissertation and to my development as a Ph.D. student. I am deeply indebted to the current directors of the IDEA program, Francesc Obiols-Homs and David Pérez-Castrillo, for putting the greatest effort in giving me all the help I needed. A special thanks goes to Ángels and Mercè for their top-notch administrative work.

I offer my heartfelt gratitude to all of my colleagues for their affection, presence, and support: I am honored to have had the chance to share my Ph.D. experience with you guys. You are really wonderful people.

A warm thanks goes to all of my friends, and all the amazing people who loved me and cared for me during these years: this thesis is also yours.

Finally, I thank my family: Matteo, Giuliana, Luca, Massimo, and Gloria. You are part of this.

Contents

Preface

This thesis comprises two separate research interests. The first one deals with understanding how risk-sharing arrangements affect household behavior in village economies. The second one refers to how trust and uncertainty about the benefits of cooperation affect network formation.

Fertilizer subsidies play a critical role in developing countries, where fertilizer use keeps lagging behind the rates recommend by agricultural experts. Which factors are restricting farmers from using the recommended amounts of fertilizer? In Chapter 1, I show the importance of risk-sharing arrangements in holding down fertilizer use in rural India and analyze how public policy can use fertilizer subsidies to fight the inefficiencies associated with these arrangements. I study a model of risk sharing in which households' choices of effort and fertilizer are private. Private information generates a moral hazard problem: risk sharing induces households to free-ride on each other's efforts. Moreover, effort provision is related to fertilizer use through a relationship of complementarity. Thus, risk sharing (which induces farmers to curtail their effort) decreases the productivity of fertilizer, ultimately leading to fertilizer being under-demanded. A fertilizer subsidy increases welfare because, by inducing farmers to buy more fertilizer, it pushes them to exert more effort, thereby weakening the bite of the moral hazard problem. I test this theory in the context of 18 villages in the Indian semi-arid tropics, with data coming from survey interviews conducted from 2009 to 2014. The effect of risk sharing on fertilizer used and hours worked is large: when going from no sharing to full insurance, average fertilizer used drops by four times and average hours worked drop by more than six times. Moreover, I show that a subsidy that would cut the observed prices of fertilizer in half would generate a consumption-equivalent gain in welfare of 51%.

Social networks play a key role in shaping many economic outcomes, such as information transmission, trade in decentralized markets, and social learning. Which factors are important in determining the relationships that end up forming? In Chapter 2, I analyze, together with Juan Camilo Cardenas, Danisz Okulicz, and Tomás Rodríguez-Barraquer, how people's trust affects the social networks they form. We measure trust for 72 members of a cohort of first-year undergraduates before they had a chance to meet and socialize. We measure people's trust using both a standard trust experiment and survey questions. After four months, we elicit five social networks among the students. Moreover, we retrieve and control for a large set of observables, including many characteristics which are likely to play a role in network formation and may be correlated with trust. We find that trust poorly explains the formation of the networks we retrieve. In particular, the effect of homophily in socio-economic background can go so far as being one order of magnitude bigger than the effect of trust. In Chapter 3, I study theoretically how uncertainty about the benefits of cooperation affects coalition formation. Two agents can agree to cooperate while holding a common prior belief about whether the other is a lemon or a peach. Each agent prefers cooperating with a peach to autarky but would stay in autarky rather than cooperating with a lemon. A utilitarian social planner can draw a noisy public signal of whether the agents are lemons before they might agree to cooperate. Drawing a signal can decrease expected welfare. Moreover, the relationship between the welfare gain of drawing a signal and the noise of the signal can have at most one discontinuity and be non-monotonic.

Chapter 1

Risk Sharing, Private Information, and the Use of Fertilizer

1.1 Introduction

In 2016, the Indian government spent about 11 billion dollars (about 0.5% of the Indian GDP) in fertilizer subsidies (Government of India, Ministry of Finance (2016)). These policies are justified by the argument that higher fertilizer use leads to increased agricultural yields, thus improving rural households' standards of living. This view is advocated by many agricultural experts, and has spurred economists' interest in uncovering which factors constrain farmers from using the recommended amounts of fertilizer.[1] In this chapter, I analyze the role of risk-sharing arrangements in holding down fertilizer use in rural India.

Rural households in low-income countries face severe income fluctuations. These households insure against idiosyncratic income shocks by relying on a variety of informal insurance (risk-sharing) arrangements, such as gift exchange or informal loans.[2] A main finding in the literature is that risk sharing does not generally reach perfect consumption smoothing (Cochrane (1991), Mace (1991), and Townsend (1994)). The presence of frictions that impede full insurance can rationalize imperfect risk sharing. A leading explanation is private effort.[3] This friction is a relevant barrier to risk sharing in many contexts.[4] The intuition is that when effort is imperfectly monitored, insurance induces

[1]Possible explanations are credit constraints and risk (Dercon & Christiaensen (2011)), the use of complementary inputs (Beaman et al. (2013)), low quality (Bold et al. (2017)), and behavioral biases (Duflo et al. (2011)).

[2]See Fafchamps (2011) for a review.

[3]Other possible frictions are limited commitment (Ligon et al. (2002)) and hidden income (Kinnan (2017)).

[4]See Ligon (1998) for evidence from rural India; Paulson et al. (2006) and Karaivanov & Townsend (2014) for evidence from rural and semi-urban Thailand; Kocherlakota & Pistaferri (2009) for evidence

households to shirk. On the other hand, the technology adoption literature has consistently shown that effort and fertilizer are complementary inputs.[5] Given that fertilizer and effort are complements, and since insurance can decrease households' incentives to exert effort, I argue that risk sharing may hold down fertilizer use through its discouraging effect on effort supply.

In this chapter, I explore the connection between risk sharing under private information in production decisions and households' incentives to use fertilizer. I offer a theoretical framework that relates the efficient level of insurance to the demand of fertilizer through households' effort supply. I show that insurance leads to inefficiently low levels of effort, thus reducing households' incentives to use fertilizer. I also prove that a fertilizer subsidy improves welfare. Empirically, I first carry out a reduced-form strategy to test the main predictions of the model. Then, I structurally estimate the model to quantify the effect of risk sharing on fertilizer use and the welfare gains from a hypothetical fertilizer subsidy.

I outline a simple model of risk sharing in which households insure themselves by sharing the profits of agricultural production. Each household can supply costly effort to its own fields and buy fertilizer to increase expected output. I characterize the constrained-efficient allocation of effort and fertilizer subject to a fixed level of insurance. Risk sharing reduces the private marginal benefit of effort, thereby inducing households to shirk. Moreover, as insurance decreases effort provision, higher risk sharing lowers the use of fertilizer as long as effort and fertilizer are complements. I then characterize the optimal level of risk sharing, and analyze the effect of a fertilizer subsidy on welfare. I show that the effect of a marginal decrease in the price of fertilizer on welfare can be decomposed in two parts. First, a fertilizer subsidy reduces the monetary costs of agricultural production, thereby increasing profits and welfare. Second, the subsidy induces households to buy more fertilizer and, because effort and fertilizer are complements, it pushes them to exert more effort. Since effort is generally under-provided in the constrained-efficient allocation, the subsidy moves the effort allocation closer to the full-information benchmark, thereby increasing welfare.

I test the model empirically using the latest (2009-2014) ICRISAT panel from rural India. First, I provide reduced-form evidence about the main predictions of the model: better insured households should be exerting less effort, and that they should be using less fertilizer as long as effort and fertilizer are complements. First, I show that insurance is negatively correlated with effort provision. More specifically, I find that, on average,

from Italy, the UK, and the USA; and Attanasio & Pavoni (2011) for evidence from the UK.

[5] See Foster & Rosenzweig (2009), Foster & Rosenzweig (2010), and Foster & Rosenzweig (2011) for evidence from India; Beaman et al. (2013) for evidence from Mali; Ricker-Gilbert (2014) for evidence from Malawi; and Haider et al. (2018) for evidence from Burkina Faso.

households with a higher average effort supply over available time periods experience higher elasticity of consumption to idiosyncratic income shocks. I then provide evidence of the complementarity between effort and fertilizer by showing that effort provision and fertilizer use are strongly positively correlated after controlling for land area, unobserved household heterogeneity, and possible seasonalities in agricultural production. Finally, as suggested by the theory, I show that insurance is negatively correlated with fertilizer use.

Then, I structurally estimate the model. I do so by fitting the theoretical relationship between the relative choices of fertilizer and effort on one hand, and the price of fertilizer, the marginal disutility of effort, and the wedge between the social and the private marginal benefit of effort on the other. A clear advantage of the structural approach is that it allows me to conduct counterfactuals and policy simulations. Moreover, it provides a joint test of (i) the relationship between risk sharing and fertilizer use, and (ii) the complementarity between effort and fertilizer. I retrieve the elasticity of substitution between effort and fertilizer, the marginal disutility of effort, and the wedge between the social and the private marginal benefit of effort. I quantitatively assess the role of risk sharing in effort supply and fertilizer use, and simulate the effects of a fertilizer price subsidy on risk sharing and welfare. I find that when going from no sharing to full insurance, effort supply decreases by more than six times and fertilizer use drops by four times. As for the fertilizer subsidy, I show that cutting the price of fertilizer in half would cause a 13% drop in risk sharing and generate a consumption-equivalent gain in welfare of 51%.

Overall, my results suggest that when there are private information frictions in agricultural production decisions, risk sharing discourages the use of inputs that complement effort, because insurance drives effort supply down. On the positive side, these results reveal that informal insurance arrangements are an important driver of fertilizer use in rural India. On the normative side, they show by how much a fertilizer subsidy can increase welfare.

Related Literature

Time and again, development economists have placed insurance among the key factors in the process of economic development. While some have argued that a lack of insurance can hold households back from adopting high-risk and high-return technologies (Rosenzweig & Binswanger (1993) and Dercon & Christiaensen (2011)), others have pointed out that pressures to share income with neighbors and relatives might reduce investment incentives (Jakiela & Ozier (2015)). This chapter contributes to the understanding of the relationship between insurance and technology adoption by analyzing a new mechanism that relates private information frictions to the complementarity between effort and fertilizer. Specifically, I argue that risk sharing decreases the use of inputs that complement effort because it discourages effort provision in the first place.

First and foremost, this chapter contributes to the understanding of the drivers of fertilizer use, focusing on the role played by risk sharing. Uncovering the determinants of agricultural input use is extremely important from both an academic and a policy perspective (Feder et al. (1985), Sunding & Zilberman (2001), Foster & Rosenzweig (2010), Udry (2010), and Jack (2013)). Indeed, most agricultural policies for poverty reduction aim to foster the use of agricultural inputs, especially fertilizer. The literature has argued before that it is critical to uncover the impact of risk-sharing arrangements on technology adoption and agricultural input use, as these arrangements are ubiquitous in village economies.[6] However, there are few papers on this topic. The only exceptions (Giné & Yang (2009) and Dercon & Christiaensen (2011)) analyze the case of new technologies, whose inherent uncertainty over benefits and costs discourages their use. In this case, better insurance should be associated with higher take-up rates. This chapter takes a very different approach. It does not speak to the issue of understanding of how insurance might boost the use of new and riskier technologies. Instead, it focuses on the discouraging effect of insurance on effort supply, and how this effect relates to the use of fertilizer through its complementarity with effort.

The mechanism I propose to link risk sharing to fertilizer use relies on a private effort friction. Most of the literature on sharecropping assumes that effort is private (Quibria & Rashid (1984), N. Singh (1991), and Sen (2016)). More importantly, private effort has been used to rationalize imperfect insurance in village economies (Ligon (1998)). While several papers have provided evidence for private effort by testing models of imperfect insurance against each other (Ligon (1998), Ábrahám & Pavoni (2005), Kaplan (2006), Attanasio & Pavoni (2011), and Karaivanov & Townsend (2014)), this friction has been considered hard to detect using observational data (Foster & Rosenzweig (2001)).[7] I contribute to this literature by providing a first *direct* evidence of a negative relationship between insurance and effort.[8] By doing so, I confirm the main implication of the private-effort explanation to imperfect insurance.

According to Foster & Rosenzweig (2010), research on agricultural input use needs to take into account the complementarity and substitutability between inputs, and in partic-

[6] According to Udry (2010), understanding "how [...] imperfect insurance influence input choice and/or technology adoption in agriculture" is "a key research agenda" in agricultural and development economics.

[7] Despite this shortcoming, there exists experimental evidence of the effect of private effort on risk sharing (Prachi (2016)).

[8] The literature on sharecropping has produced consistent evidence that better risk sharing (in the form of a lower fraction of the agricultural output going to the tenant) leads to lower efficiency and effort provision (Laffont & Matoussi (1995)). However, the same empirical evidence has not been provided by the literature on informal insurance.

ular between labor and agricultural intermediates. Indeed, empirical evidence (Dorfman (1996) and Hornbeck & Naidu (2014)) suggests that labor availability plays a key role in the decision to adopt different input baskets. By taking into account the complementarity between effort and fertilizer, my model directly speaks to this issue. In particular, the model explicitly recognizes that the profitability of an agricultural input (and hence its use) will ultimately depend on a household's willingness to allocate its time to farm labor (which is in turn affected by how insured it is).

Finally, this chapter relates to a growing literature focusing on how informal insurance affects different aspects of the village economy. Important contributions to this literature are Munshi & Rosenzweig (2006), which studies how risk sharing shapes career choice by gender in Bombay; Munshi & Rosenzweig (2016), which analyzes how caste-based informal insurance affects incentives to migrate in India; Advani (2017), which studies how informal insurance with limited commitment impacts on investment in livestock in Bangladesh; and Morten (2017), which studies the joint determination of informal insurance and temporary migration in rural India when there is a limited commitment friction.

1.2 A Model of Risk Sharing

I study an economy in which households face productivity shocks, and insure themselves by pooling part of their incomes together and sharing them equally. Each household chooses how much effort to supply to its own agricultural business and how much fertilizer to buy at a given price. In Subsection 1.2.1 I outline the setup of the model. In Subsection 1.2.2 I characterize the efficient allocation of effort and fertilizer as a function of the sharing contract, both when the households' choices are public and when they are private. In Subsection 1.2.3 I solve for the efficient sharing rule in both the full- and private-information regimes. Finally, in Subsection 1.2.4 I discuss the effect of a fertilizer subsidy on welfare. All the proofs are contained in Appendix A.9.

1.2.1 Setup

There are n household-farms, each producing output y_i, $i \in N := \{1, \dots, n\}$. Output is uncertain, and depends on effort $e_i \in [0, \overline{e}_i]$ and fertilizer $f_i \in \mathbb{R}_+$. Refer to $a_i := (e_i, f_i)$ as an action for household i. Let ε_i be a production shock with mean μ and variance η^2. I assume that

$$y_i := y\left(a_i\right) + \varepsilon_i, \tag{1.1}$$

where y is jointly concave in a_i, and strictly concave, strictly increasing, and twice-continuously differentiable in both e_i and f_i. Hence, supplying more effort or increasing the use of fertilizer increases expected output without making it riskier.[9] The shocks are independently distributed.[10] Each household can only supply effort to its own agricultural business (i.e., there is no market for effort). On the other hand, fertilizer is bought in the market, and households take its price as given. Letting $p \in \mathbb{R}_{++}$ be the price of fertilizer, household i's agricultural profit (income) is given by

$$\pi_i := y_i - pf_i. \tag{1.2}$$

Households share incomes[11] to smooth consumption risk. In particular, household i's consumption is given by

$$c_i(\alpha) := (1-\alpha)\pi_i + \alpha\overline{\pi}, \tag{1.3}$$

where $\alpha \in [0,1]$ is a coefficient that fully characterizes the extent of risk sharing and $\overline{\pi}$ is average income. The intuition is that each household consumes a fraction $1-\alpha$ of its agricultural profit, and contributes the rest to a common pool that is shared equally.[12] Risk sharing is enforceable.

Household i's expected utility is given by[13]

$$U(c_i(\alpha), e_i) := \mathbb{E}(c_i(\alpha)) - \frac{\rho}{2}\mathbb{V}\text{ar}(c_i(\alpha)) - \kappa e_i,$$

[9] See Appendix A.2 for a discussion of how the results here presented can be preserved in a model where effort and fertilizer have an impact on the variance of output. Fertilizer is generally found to be a risk-increasing input (Just & Pope (1979)). If the impact of fertilizer on yield variability is not too big, and either effort makes production less risky or it does not increase output volatility by too much, then my qualitative results are preserved.

[10] This assumption is without loss of generality as long as the shocks are not perfectly correlated. Indeed, one can assume that $\varepsilon_i := v + \theta_i$, where v is an aggregate shock and θ_i is idiosyncratic risk. In this case, it is optimal for the households to only share the idiosyncratic component of their shocks.

[11] Since I assume that income equals agricultural profit, households share both outputs and the monetary costs of fertilizer. This assumption simplifies the analysis of the model, as it implies that risk sharing only has a direct impact on effort choices. Of course, in equilibrium risk sharing does affect fertilizer use, but this effect only comes about through its impact on effort. The informal insurance literature made it customary to think of households as sharing outputs instead of profits. Appendix A.3 shows that my qualitative results are preserved when households share outputs instead of profits.

[12] Equation (1.3) can be thought of as an implementation of the well-known contrast estimator (Suri (2011)) when the economy is closed and there are no saving technologies, as shown in Appendix A.4.

[13] The assumption that expected utility is separable in consumption and effort is standard in the moral hazard literature. I make use of a mean-variance expected utility of consumption because it greatly simplifies strategic interactions between households: given some α, i's choices of effort and fertilizer do not depend on other households' choices. See Appendix A.1.1 for a more detailed discussion. Also notice that the disutility of effort is linear in effort, so that the marginal disutility of effort is just another price. This modeling strategy allows me to apply standard results in producer theory in what follows. The same strategy has been widely used in papers dealing with sharecropping (Arcand et al. (2007)) and general agency problems (Conlon (2009)).

where ρ is the coefficient of absolute risk aversion and κ is the marginal disutility of effort. Expectations are taken with respect to the distribution of the production shocks.

For simplicity, price p can be thought of as a parameter, thus making this model a partial equilibrium model. Equivalently, one can assume that fertilizer is supplied by competitive manufacturers with a linear cost function cf, for some $c > 0$. In this case, profit maximization on the part of the manufactures implies that $p = c$, which pins down the equilibrium price of fertilizer. Since there are no markets for effort and consumption, this 'price equals marginal cost' condition characterizes the unique (general) equilibrium for the economy.

An allocation is a sharing rule α together with an action profile $\boldsymbol{a} := (a_i)_i$. There is a utilitarian social planner who chooses an allocation to maximize welfare. My aim is to characterize a welfare-maximizing allocation in two information regimes: full information and private information. A welfare-maximizing allocation under full information is said to be efficient, while a welfare-maximizing allocation under private information is said to be constrained efficient. In order to solve the planner's problem, I proceed as follows. First, I find a welfare-maximizing action profile $\boldsymbol{a}^*$ for a given sharing rule α. Then, I find a welfare-maximizing sharing rule α^*.

1.2.2 Optimal Action Profile

Full information. Assume that the planner observes $\boldsymbol{a}$. The problem of finding a welfare-maximizing action profile for a given α is

$$\max_{\boldsymbol{a}} \sum_{i \in N} U\left(c_i\left(\alpha\right), e_i\right), \tag{1.4}$$

subject to Equations (1.3), (1.2), and (1.1). Notice that there are no participation constraints. This is without loss of generality, as the planner is benevolent and each household's Pareto weight is 1. Let $\boldsymbol{a}^{\diamond}\left(\alpha\right)$ be a solution to Problem (1.4). The following claim nails down the welfare-maximizing action profile.

Claim 1.2.1 (Efficient action profile)**.** *Under full information, and for given α, the welfare-maximizing action profile is characterized by*

$$y_e\left(a_i^{\diamond}\left(\alpha\right)\right) = \kappa,$$
$$y_f\left(a_i^{\diamond}\left(\alpha\right)\right) = p.$$

The intuition behind this claim is straightforward: under full information risk sharing does not generate externalities; hence, the optimal action profile is independent of α. In particular, the planner equates the marginal product of effort to its marginal utility cost, and the marginal product of fertilizer to its market price.

Private information. Next, assume that household i's action is private to i. In this case, to find a welfare-maximizing action profile for a given α the planner has to solve

$$\begin{aligned} &\max_{\boldsymbol{a}} \sum_{i \in N} U\left(c_i\left(\alpha\right), e_i\right), \\ &\text{subject to } a_i \in \arg\max_{\widehat{a}_i} U\left(c_i\left(\alpha\right), \widehat{e}_i\right), \ \forall i \in N, \end{aligned} \tag{1.5}$$

and Equations (1.3), (1.2), and (1.1). The difference between Problems (1.4) and (1.5) is that in the private information regime an optimal action profile has to satisfy n incentive-compatibility (IC) constraints, which say that the action chosen by the planner for household i coincides with what the household would do on its own; otherwise, the household has an incentive to deviate to another action. Let $\boldsymbol{a}^*(\alpha)$ be a solution to Problem (1.5). The following claim characterizes the solution to Problem (1.5).

Claim 1.2.2 (Constrained-efficient action profile)**.** *Under private information, and for given α, the welfare-maximizing action profile is characterized by*

$$\begin{aligned} y_e\left(a_i^*\left(\alpha\right)\right) &= \frac{\kappa}{\left(1 - \frac{n-1}{n}\alpha\right)} =: p_e, \\ y_f\left(a_i^*\left(\alpha\right)\right) &= p, \end{aligned}$$

for each $i \in N$.

Refer to p_e as the 'effective price' of effort. Claim 1.2.2 shows that risk sharing induces a direct negative externality on effort provision, as it increases the 'effective price' of effort. On the other hand, risk sharing has no direct impact on fertilizer use, because it affects neither its marginal benefit nor its marginal cost. This asymmetry between effort and fertilizer comes from the assumptions that (i) households share profits, so that they share both the revenues and the monetary costs of production, and (ii) effort does not enter the monetary costs of production, as there is no market for effort. The impact of the sharing contract on the private marginal benefit of fertilizer cancels out with its impact on its private marginal cost. This is not the case for effort: the sharing contract decreases its private marginal benefit while leaving unaltered its private marginal cost. The next theorem shows how effort supply and fertilizer use change when the sharing coefficient moves.

Theorem 1.2.1 (Effort, fertilizer, and risk sharing)**.** *Let $\boldsymbol{a}^*(\alpha)$ be the constrained-efficient action profile. Then,*

$$\frac{\partial e_i^*(\alpha)}{\partial \alpha} < 0.$$

Moreover, suppose that e_i and f_i are complements at (p_e, p); i.e.,

$$\frac{\partial f_i^*(\alpha)}{\partial p_e} < 0.$$

Then,

$$\frac{\partial f_i^*(\alpha)}{\partial \alpha} < 0.$$

The signs of the latter two inequalities are reversed if e_i and f_i are substitutes at (p_e, p).

Theorem 1.2.1 shows that if risk sharing increases, then households exert less effort, and decrease the use of fertilizer as long as effort and fertilizer are complements.[14] The intuition is as follows. Since effort has a direct negative externality on effort provision, more insurance induces households to shirk. This reduction in effort induces the households to decrease fertilizer use, as it becomes less profitable.

1.2.3 Optimal Sharing Rule

In this model sharing contracts are assumed to be linear.[15] Nevertheless, the results obtained in this section carry through to more complex environments in which the linearity assumption is dropped, as shown in Appendix A.1.

Full information. Next, consider the problem of finding a welfare-maximizing sharing contract under full information; i.e.:

$$\max_{\alpha} \sum_{i \in N} U\left(c_i(\alpha), e_i\right),$$

subject to Equations (1.3), (1.2), (1.1), and $\boldsymbol{a} = \boldsymbol{a}^{\diamond}(\alpha) =: \boldsymbol{a}^{\diamond}$, where $\boldsymbol{a}^{\diamond}$ is the solution to Problem 1.4. The following claim shows that, under full information, risk sharing is perfect.

[14]See Appendix A.5 for a discussion on the definition of complementarity and substitutability used in Theorem 1.2.1, as well as Theorem A.1.1 in Section A.1, and other possible definitions of complementarity and substitutability (and additional assumptions on the production function) under which these theorems would still hold.

[15]In general, linear contracts are not optimal when there is a private information friction. Yet, linearity simplifies the analysis considerably. Moreover, linear contracts can be motivated by empirical evidence, as in J. Dutta & Prasad (2002). In fact, explaining why linear contracts are so common is a longstanding problem in contract theory, since most models predict more complicated functional forms (Holmström & Milgrom (1987) and Carroll (2015)).

Claim 1.2.3 (Efficient sharing)**.** *Under full information, the welfare-maximizing sharing contract is full insurance.*

This result is entirely standard: since under full information risk sharing does not generate externalities, the planner maximizes welfare by insuring the households as much as possible.

Private information. Finally, assume that the households' choices are private. In this case, the problem of finding a welfare-maximizing sharing contract is given by

$$\max_{\alpha} \sum_{i \in N} U\left(c_i\left(\alpha\right), e_i\right),$$

subject to Equations (1.3), (1.2), (1.1), and $\boldsymbol{a} = \boldsymbol{a}^*\left(\alpha\right)$, where $\boldsymbol{a}^*\left(\alpha\right)$ is the solution to Problem 1.5. Let $W\left(\alpha\right)$ denote social welfare evaluated at $\boldsymbol{a}^*\left(\alpha\right)$. The next claim characterizes the welfare-maximizing sharing contract under private information, and highlights that, under this information regime, a marginal increase in α generates a trade-off between decreasing consumption volatility and decreasing aggregate consumption.

Claim 1.2.4 (Constrained-efficient sharing)**.** *First, notice that*

$$\frac{\partial W\left(\alpha\right)}{\partial \alpha} = \underbrace{\sum_{i \in N} \left(\kappa \left(\frac{1}{1-\frac{n-1}{n}\alpha} - 1\right) \frac{\partial e_i^*\left(\alpha\right)}{\partial \alpha}\right)}_{(-)} \underbrace{- \frac{n\rho}{2} \frac{\partial \mathbb{V}\mathrm{ar}\left(c_i\left(\alpha\right)\right)}{\partial \alpha}}_{(+)}. \tag{1.6}$$

Let α^* be an optimal sharing rule under private information. It must be the case that

$$\begin{aligned} \tfrac{\partial W(\alpha^*)}{\partial \alpha} &= 0 \quad \textit{if } \alpha^* \in (0,1)\,, \\ \tfrac{\partial W(\alpha^*)}{\partial \alpha} &\leq 0 \quad \textit{if } \alpha^* = 0, \\ \tfrac{\partial W(\alpha^*)}{\partial \alpha} &\geq 0 \quad \textit{if } \alpha^* = 1. \end{aligned}$$

The first term of Equation (1.6), which is negative, is the loss in aggregate production generated by a marginal increase in the negative externality caused by sharing. This is the marginal cost of risk sharing. The second term, which is positive, is the gain associated with a marginal reduction in consumption volatility. This is the marginal benefit of risk sharing. An optimal sharing rule balances the trade off between effort provision and consumption smoothing. Hence, in general, one should not expect to observe $\alpha = 1$, as under full information.

1.2.4 Fertilizer Subsidy

Next, I analyze the effect of a fertilizer subsidy on welfare. First notice that welfare can be written as

$$\sum_{i\in N}\left[y\left(a_i\right)-pf_i-\kappa e_i-\frac{\rho}{2}\left(\left(1-\alpha\right)^2+\frac{\alpha^2}{n}+\frac{2\alpha\left(1-\alpha\right)}{n}\right)\eta^2\right].$$

Exogenous sharing rule. First, assume that α is fixed. By the envelope theorem, the effect of a marginal *decrease* in the price of fertilizer on welfare under full information is given by

$$\sum_{i\in N} f_i^{\diamond}.$$

The subsidy increases profits by mechanically reducing the monetary costs of agricultural production. I call this the price effect.

Under private information, the effect of a marginal decrease in the price of fertilizer on welfare is

$$\sum_{i\in N}\left[-\left(y_e\left(a_i^*\left(\alpha\right)\right)-\kappa\right)\frac{\partial e_i^*\left(\alpha\right)}{\partial p}+f_i^*\left(\alpha\right)\right].$$

The subsidy does not only reduce the monetary costs of production but it also affects effort supply. Recall that $y_e\left(a_i^*\left(\alpha\right)\right)-\kappa>0$ (see Claim 1.2.2): since effort is under-provided under private information, its marginal product is greater than its marginal cost. When fertilizer and effort are complements (which implies $\partial e_i^*\left(\alpha\right)/\partial p<0$), the subsidy induces households to exert more effort, thus shrinking the negative externality generated by risk sharing. I call this the *direct* effort effect.

Endogenous sharing rule. Next, assume that insurance is endogenous. Under full information, the welfare-maximizing sharing rule is full insurance irrespective of the price of fertilizer (see Claim 1.2.3). Thus, the overall effect of the subsidy coincides with the price effect. On the other hand, under private information, insurance responds to changes in the price of fertilizer. This is because, by affecting the households' incentives to exert effort, the subsidy affects the marginal cost of risk sharing; i.e., the reduction in effort supply given rise by a marginal increase in insurance. In order to determine how insurance responds to the subsidy, recall from Claim 1.6 that an interior optimal sharing rule is implicitly defined by

$$\frac{\partial W\left(\alpha^*\right)}{\partial\alpha}=0.$$

By the implicit function theorem, the effect of a marginal decrease in the price of fertilizer on optimal insurance is given by

$$-\frac{\partial \alpha^*}{\partial p} = \frac{\frac{\partial^2 W(\alpha^*)}{\partial \alpha \partial p}}{\frac{\partial^2 W(\alpha^*)}{\partial \alpha^2}}.$$

Notice that local optimality implies that $\partial^2 W(\alpha^*) / \partial \alpha^2 < 0$. Moreover,

$$\frac{\partial^2 W(\alpha^*)}{\partial \alpha \partial p} = \sum_{i \in N} \left[\kappa \left(\frac{1}{1 - \frac{n-1}{n}\alpha} - 1 \right) \frac{\partial^2 e_i^*(\alpha^*)}{\partial \alpha \partial p} \right].$$

Hence,

- if $\partial^2 e_i^*(\alpha^*) / \partial \alpha \partial p > 0$, the subsidy decreases insurance;
- if $\partial^2 e_i^*(\alpha^*) / \partial \alpha \partial p = 0$, the subsidy does not affect insurance;
- if $\partial^2 e_i^*(\alpha^*) / \partial \alpha \partial p < 0$, the subsidy increases insurance.

To gain intuition, notice that $\partial e_i^*(\alpha^*) / \partial \alpha$ is the decrease in effort supply associated to a marginal increase in the sharing rule; i.e., the slope of the effort supply function with respect to risk sharing. This is the marginal cost of insurance: the more negative this slope, the more costly is insurance in terms of reducing effort provision. Recall that the marginal benefit of insurance (i.e., the marginal increase in consumption smoothing) is independent of the price of fertilizer (see Equation (1.6)). If $\partial^2 e_i^*(\alpha^*) / \partial \alpha \partial p > 0$ then the slope of the effort supply function with respect to risk sharing becomes more negative when the price of fertilizer is lower. Hence, a fertilizer subsidy increases the marginal cost of insurance, making it bigger than its marginal benefit. Because of the concavity of the welfare function around α^*, the planner decreases α to reestablish the equality between the marginal benefit and the marginal cost of risk sharing.

The change in risk sharing generated by the subsidy has two effects on welfare in two ways, but the sum of the two effects is zero. On one hand, as shown in Theorem 1.2.1, a decrease in risk sharing leads households to exert more effort, while an increase in risk sharing induces households to shirk more.[16] From this point of view, a decrease in insurance is good for welfare because it moves the effort allocation closer to the full-information benchmark. I call this the *indirect* effort effect. On the other hand, a decrease in risk sharing increases consumption volatility, while an increase in risk sharing reduces consumption volatility. From this point of view, a decrease in insurance is bad for welfare, because it makes consumption more responsive to idiosyncratic shocks. I call this the

[16] The change in risk sharing also affects fertilizer use. However, because of the envelope theorem, the welfare effect of this change in fertilizer use is null.

consumption smoothing effect. The insurance effect is given by the sum of the indirect effort effect and the consumption smoothing effect. Because of the envelope theorem, the sum of the indirect effort effect and the consumption smoothing effect is zero. Hence, the overall effect of the subsidy on welfare is given by the sum of the the price effect and the direct effort effect, which are both welfare-enhancing. Hence, a fertilizer subsidy increases welfare.

1.3 Empirical Evidence

This section presents a description of the data used, reduced-form evidence confirming the main theoretical predictions linking risk sharing and fertilizer, and a structural estimation of the model outlined above. Equipped with the the estimated structural parameters, I conduct a counterfactual exercise in which I quantify the impact of risk sharing on fertilizer use, and calculate the welfare gains from a fertilizer subsidy.

1.3.1 Background and Data

I use data collected under the Village Dynamics in South Asia (VDSA) project by the International Crop Research Institute for the Semi-Arid Tropics (ICRISAT). This data-set is a household-level panel covering more than 700 households in 18 villages in the Indian semi-arid tropics. The data come from survey interviews conducted at an approximately monthly frequency from 2009 to 2014. For each village, 40 households were randomly selected stratifying by landholding classes: 10 from landless laborers, 10 from small farmers, 10 from medium farmers, and 10 from large farmers. I choose this data for two reasons: first because it provides detailed information on farming activity, expenditure, and income; and second because it has been widely used to test risk sharing models (hence, my results can be compared with the findings of earlier papers). I refer to Townsend (1994), Mazzocco & Saini (2012), and Morten (2017) for more detailed descriptions of the data.

There are three problems with the data: (i) the frequency of the interviews varies, (ii) the interview dates differ across households, and (iii) recall periods vary across interviews. Fortunately, from 2010 onwards, information is provided on the month and year to which a given interview refers to. Since recall periods can be longer than a month, it is impossible to determine to which month an interview refers to if this information is not provided. Therefore, I drop the observations from 2009.

For the estimation, I need information on demographic characteristics, consumption, income, fertilizer, a proxy for effort, and the price of fertilizer. In the following, I discuss how I construct these variables. I convert money values to 1975 rupees for comparability with Townsend (1994).

Following Mazzocco & Saini (2012), I use the data coming from the General Endowment Schedule to build a set of observable household heterogeneity variables, which I use to build the age-sex weight proposed in Townsend (1994).

Monthly household consumption is calculated using the Transaction Schedule. This schedule reports household-level data on the value of items purchased, home produced, and acquired in other ways (such as gifts). Following Kinnan (2017), I sum the value of all items across categories to build a measure of total household expenditure. Since different households have different sizes and age-sex structure, I convert total household expenditure to adult-equivalent terms by dividing it by the age-sex weight.

Monthly household income is calculated following Mazzocco & Saini (2012)'s methodology. Making use of the household budget constraint, I compute total income as total expenditure minus resources borrowed, plus resources lent and saved, minus government benefits. The Transaction Schedule contains information on these variables. I aggregate the data following the same procedure I use to calculate monthly household consumption. Again, I convert total household income to adult-equivalent terms using the age-sex weight.

The Cultivation Schedule contains information on the quantity and total value of each agricultural intermediate and the labor used in all the operations performed (e.g., harvesting, irrigating, sowing) on every plot cultivated by the household. A distinction is made between family, hired, and exchanged labor. To build a proxy for monthly effort supply, I first compute the total amount of hours of work supplied by family members to the operations performed on all the plots cultivated by the family in each month. Then, I convert this measure to adult-equivalent terms using the age-sex weight.[17]

The Cultivation Schedule classifies agricultural intermediates by their names (e.g., DAP, Pursuit, and Acephate) and the type of operation in which they are used (e.g., fertilizer application and plant protection), but does not give information on their type (e.g., fertilizer, herbicide, and pesticide). Retrieving the type of each agricultural intermediate from its name is a time-consuming process, as there are almost 1000 intermediate names and due to numerous spelling mistakes. My focus is on fertilizers, by which I mean any substance that supplies plant nutrients or amend soil fertility.[18] To aggregate use of different fertilizers into a per capita measure of fertilizer use at the household-month level, I sum the monthly *value* (instead of the physical quantity) of each intermediate

[17]Proxing effort by hours of work is not uncommon: see e.g. Clark et al. (2003).

[18]Fertilizers are applied through soil for uptake by plant roots, or by applying liquid fertilizer directly to plant leaves. Chemical fertilizers are named according to the content of three macronutrients: nitrogen, phosphorus, and potassium.

used in the fertilizer application operations performed on every plot (see C. Chen et al. (2017)). Then, I convert this measure to adult-equivalent terms using the age-sex weight. The Cultivation Schedule reports the value of fertilizer on the basis of the prevailing market price (see R. P. Singh et al. (1985)). For my purposes, the key is that per capita fertilizer use at the household-month level reflects physical variation in physical quantity of different types of fertilizers. Valuing output at common market prices therefore allows me to compare per capita fertilizer use across households and months, reflecting variation in quantity of fertilizers used.

I calculate fertilizer prices using the Monthly Price Schedule. Interviews on fertilizer prices are not conducted at the household level. Instead, in each village, five respondents—labeled A, B, C, D, and E—are asked to report the average monthly price of each fertilizer. These respondents correspond to a large farmer, a medium farmer, a landless laborer, a village trader, and a trader from the nearest market.[19] Hence, for each village-month pair, five different prices are reported for each fertilizer. Since my fertilizer use variable is defined at the household-month level, I need to aggregate these fertilizer prices at the household-month level. For each respondent, I average fertilizer prices across types of fertilizers. Then, I assign the average price reported by the large farmer (respondent A) to large landholding households and the average price reported by the medium farmer (respondent B) to medium landholding households. Finally, I average the average fertilizer prices across all respondents and assign this average price to the rest of the households.

Summary statistics for the sample are reported in Table A1.

1.3.2 Reduced-Form Evidence Linking Agricultural Production and Risk Sharing

I document the following facts in the data: (i) risk sharing is imperfect; (ii) risk sharing and effort supply are negatively correlated; (iii) effort supply and fertilizer use are positively correlated; and (iv) risk sharing and fertilizer use are negatively correlated.

Risk sharing is imperfect. I estimate the following regression for household i in village v and month t:

$$\log(c_{it}) = \beta_1 \log(\pi_{it}) + \varphi_i + \phi_{vt} + \epsilon_{it}, \tag{1.7}$$

[19] In the data, I observe sizable dispersion in fertilizer prices across households and time. This might be due to costly information (Aker (2010)) or price discrimination operated by local traders.

where φ_i and ϕ_{vt} are household and village-month fixed effects. Equation (1.7) estimates the elasticity of consumption with respect to income after controlling for aggregate income shocks through village-month fixed effects. Standard errors are robust.

Table A2 reports the results of the test. Full sharing is strongly rejected. The elasticity of consumption with respect to income is about 0.26. Although the magnitude of this coefficient varies across studies, a value of 0.26 falls well within the expected range. For instance, Munshi & Rosenzweig (2009) estimate values between 0.17 and 0.26 for rural India, using data from the Rural Economic and Demographic Survey (REDS); Cochrane (1991) finds values between 0.1 and 0.2 for the United States using data from the Panel Study of Income Dynamic (PSID); Milán (2016) finds a value of 0.35 for indigenous villages in the Bolivian Amazon. Overall, the results square fairly well with the literature and unequivocally reject full insurance.

As robustness checks, I estimate Equation (1.7) aggregating the data at a quarterly and annual frequency, and run the alternative specifications outlined in Jalan & Ravallion (1999) and Mazzocco & Saini (2012). Reassuringly, the results do not change.

Effort is lower when risk sharing is higher. Theorem 1.2.1 says that effort decreases when risk sharing increases. To analyze the correlation between risk sharing and effort, I follow Morten (2017) and estimate the following regression:

$$\log(c_{it}) = \beta_1 \log(\pi_{it}) + \beta_2 \log(\pi_{it}) \overline{\log(e_i)} + \varphi_i + \phi_{vt} + \epsilon_{it}, \tag{1.8}$$

where e_{it} is the adult-equivalent total work hours supplied to own fields by household i in month t, and $\overline{\log(e_i)}$ is the average effort supplied by household i.[20] Coefficient β_2 is the correlation between average effort supplied at the household level and the elasticity of consumption with respect to the idiosyncratic component of income. If insurance is negatively correlated to effort, β_2 should be positive. In this case, the slope of consumption on income is higher as average effort increases. That is, comparing two households which are identical across any dimension captured by the household and village-month fixed effects, a positive β_2 indicates that the consumption of the household that supplied more effort is expected to be more responsive to idiosyncratic shocks to own income. Table A3 reports the results of the OLS estimation of Equation (1.8). The interaction term is positive and significant. To get a sense of the magnitudes, assume that effort supply is constant in time. Then, on average, doubling effort provision is associated to

[20]The main effect of $\overline{\log(e_i)}$ is omitted because it is captured by the household fixed effects. If I were to interact $\log(\pi_{it})$ with $\log(e_{it})$, I would need to control for the main effect of $\log(e_{it})$. Unfortunately, this strategy does not produce good results, as $\log(\pi_{it}) \log(e_{it})$ and $\log(e_{it})$ are almost collinear. This makes the standard errors explode, while barely affecting the point estimate of the interaction term.

14% increase in the elasticity of consumption with respect to income. This confirms that households that are less well insured put more effort.

Notice that, in Equation (1.8), I interact $\log(\pi_{it})$ with $\overline{\log(e_i)}$ instead of $\overline{e}_i$, hence effectively giving less weight to higher values of $\overline{e}_i$. Indeed, the relationship between insurance and effort is non-linear. This can be seen by running the following regression:

$$\log(c_{it}) = \beta_1 \log(\pi_{it}) + \beta_2 \log(\pi_{it})\,\overline{e}_i + \beta_3 \log(\pi_{it})\,\overline{e}_i^2 + \varphi_i + \phi_{vt} + \epsilon_{it}. \qquad (1.9)$$

Table A4 gives the results of the OLS estimation of the Equation (1.9). The interaction between the log of income and average effort is positive and significant, confirming the results obtained in Table A3. However, the interaction between the log of income and average effort squared is negative and significant. This suggests that as average effort supply increases, the negative relationship between effort and insurance becomes weaker. When omitting $\log(\pi_{it})\,\overline{e}_i^2$ from Equation (1.9), β_2 narrowly becomes insignificant, suggesting that the non-linearity between effort and insurance is important.

Even though these tests do not speak to causality, they are consistent with the model and provide suggestive evidence about the disincentive effect of insurance.

Effort and fertilizer. Next, I provide evidence about the complementary between effort and fertilizer. I run the following regression equation:

$$\log(f_{it}) = \gamma \log(e_{it}) + \varphi_i + \ell_{it} + \tau_t + \epsilon_{it}, \qquad (1.10)$$

where f_{it} is the adult-equivalent value of fertilizer used by household i, ℓ_{it} is land area, and τ_t are month fixed effects. My measure of fertilizer includes organic compounds (such as urea), micro-nutrients, and manure. The results are reported in Table A5. Effort is significantly positively correlated with fertilizer, which suggests the existence of a complementarity.[21] Indeed, an argument to see why it makes sense to think of fertilizer as a complement to effort is the following. Fertilizer is generally considered to be a land-augmenting technologies: in efficiency units, a quantity of fertilized land can be conceived as a multiple of a smaller quantity of unfertilized land. Hence, when land and effort are complementary, so are effort and fertilizer.

I run a version of Equation (1.10) in levels as robustness check, and find that all the results go through.[22]

[21] A formal test for the complementarity between effort and fertilizer requires one to estimate the elasticity of substitution between effort and fertilizer. I carry out this test in Subsection 1.3.3.

[22] As a cautionary note, regression (1.10) surely cannot be interpreted in a casual fashion. In particular, there is an obvious problem of reverse causality. This is why the structural estimation of the elasticity of substitution between effort and fertilizer in Subsection 1.3.3 is needed to confirm the suggestive evidence

Fertilizer is lower when risk sharing is higher. Theorem 1.2.1 implies that (i) if an input is complementary to effort, households use less of it as long as they are better insured; (ii) if an input substitutes effort, households use more of it as long as they are better insured. Before, I provided suggestive evidence about the complementarity between effort and fertilizer. Hence, I expect to observe a negative correlation between insurance and fertilizer use. To test this hypothesis, I begin by running the following regression:

$$\log(c_{it}) = \beta_1 \log(\pi_{it}) + \beta_3 \log(\pi_{it}) \overline{\log(f_i)} + \varphi_i + \phi_{vt} + \epsilon_{it}. \tag{1.11}$$

The correlation between average fertilizer use and the elasticity of consumption with respect to income is given by β_3. If insurance is negatively correlated to the use of fertilizer, β_3 should be positive, and negative otherwise. Table A6 reports the results of running regression (1.11). Indeed, the results show that β_3 is positive and significant.

For completeness, I test the non-linearity between risk sharing and fertilizer use by running the following regression:

$$\log(c_{it}) = \beta_1 \log(\pi_{it}) + \beta_2 \log(\pi_{it}) \overline{f}_i + \beta_3 \log(\pi_{it}) \overline{f}_i^2 + \varphi_i + \phi_{vt} + \epsilon_{it}. \tag{1.12}$$

Table A7 gives the results of the OLS estimation of the Equation (1.9). While coefficients β_2 and β_3 narrowly lose significance, the signs of the coefficients clearly confirm the same intuition provided by Table A4.

1.3.3 Structural Estimation of a Simple Model of Risk Sharing

I now estimate the model outlined in Section 1.2. I use the estimates obtained in this subsection to conduct a counterfactual exercise and a policy simulations. My strategy is to estimate the relative demand of fertilizer and effort, making use of Claim 1.2.2. Once I estimate this relative demand, I proceed to (i) quantify the impact of risk sharing on effort supply and fertilizer use, (ii) back out the optimal sharing rule, making use of Claim 1.2.4, and (iii) simulate the effect of a fertilizer subsidy on risk sharing and welfare.

This subsection begins by describing the identification and estimation of the model. The key advantage of this model is that strategic interactions between households are greatly simplified by the assumptions that (i) households have mean-variance expected utility and (ii) the sharing contract is linear (see Subsection A.1.1). Relaxing these assumptions would typically give rise to more convoluted strategic interactions, hence making identification and estimation substantially more complex. On the negative side, these assumptions are harmful to the richness of the model and its ability to capture

here reported.

relevant sources of variation in the data.

First, I impose a specific functional form to the production function. In particular, assume that the value of agricultural output is given by

$$y\left(a_i\right) = \ell_i^{1-\chi}\left[e_i^{\frac{\sigma-1}{\sigma}} + f_i^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\chi\sigma}{\sigma-1}},$$

where $\chi \in (0,1]$, $\sigma \in (0,\infty)$ is the elasticity of substitution between effort and fertilizer, and ℓ_i is land, which I assume to be fixed.[23] Hence, $1-\chi$ can be interpreted as the land share. With this production function (and denoting by e_i^* and f_i^* the optimal choices of effort and fertilizer for the sake of readability), the first-order conditions for effort and fertilizer given in Claim 1.2.2 read as follows:

$$\ell_i^{1-\chi}\chi\left[e_i^{*\frac{\sigma-1}{\sigma}} + f_i^{*\frac{\sigma-1}{\sigma}}\right]^{\frac{\chi\sigma}{\sigma-1}-1} e_i^{*-\frac{1}{\sigma}} = \frac{\kappa}{\left(1-\frac{n-1}{n}\alpha\right)}$$

and

$$\ell_i^{1-\chi}\chi\left[e_i^{*\frac{\sigma-1}{\sigma}} + f_i^{*\frac{\sigma-1}{\sigma}}\right]^{\frac{\chi\sigma}{\sigma-1}-1} f_i^{*-\frac{1}{\sigma}} = p.$$

Dividing the second equation by the first one, rearranging, and taking logs, I obtain

$$\log\left(\frac{f_i^*}{e_i^*}\right) = \sigma\log\left(\kappa\right) - \sigma\log\left(1-\frac{n-1}{n}\alpha\right) - \sigma\log\left(p\right).$$

In the data, I observe effort provision, fertilizer use, and the price of fertilizer. The parameters of interest are the elasticity of substitution between effort and fertilizer (σ), the marginal disutility of effort (κ), and the wedge between the social and the private marginal benefit of effort ($1-\frac{n-1}{n}\alpha$). Assuming that the model is correctly specified, if there is an error in the measurement of fertilizer or effort, I can estimate

$$\log\left(\frac{f_{it}}{e_{it}}\right) = \sigma\log\left(\kappa_i\right) - \sigma\log\left(1-\frac{n_{vt}-1}{n_{vt}}\alpha_{vt}\right) - \sigma\log\left(p_{it}\right) + \epsilon_{it}, \tag{1.13}$$

where I am assuming that the marginal disutility of effort is constant in time but possibly

[23] This production function exhibits non-increasing returns to scale in a_i. I do not need to assume decreasing returns to scale in a_i (i.e., I can assume $\chi = 1$) to back out the structural parameters, perform the counterfactual exercise, and calculate the welfare gains from a fertilizer subsidy. However, I need to assume decreasing returns in a_i (i.e., $\chi < 1$) to compute the optimal sharing rule. This is because computing the optimal sharing rule requires to calculate the responsiveness of effort supply and fertilizer use to α and the p. To see why I need decreasing returns to calculate this responsiveness, recall that the household's problem is equivalent to that of a competitive firm facing a real price of fertilizer equal to p and a real price of effort equal to $\kappa\left(1+\frac{n-1}{n}\alpha\right)^{-1}$. (See the proof of Theorem 1.2.1.) Under constant returns, the profit-maximizing choices of inputs by a competitive firm are indeterminate. So, I can either impose an additional constraint to pin down some *ad hoc* production level to back out a_i^*, or assume decreasing returns in a_i.

heterogeneous across households, and that village size and risk sharing are possibly time varying and village specific. In principle, α_{vt} could also be defined at the household level. For example, when using the time series estimation proposed by Townsend (1994), the risk sharing coefficient is assumed to be household specific and time invariant. On the other hand, if one estimates α by following a pooling strategy, as I do in Subsection 1.3.2, then this coefficient is assumed to be constant across households, villages, and time. I do not take a stance on whether α_{vt} varies across time or villages, but I do require α_{vt} *not* to be household specific. This is a necessary condition to identify the parameters of interest, as explained below.

Estimation. Under the premise that the model is correctly specified, the underlying assumptions for the consistent estimation of σ, κ_i, and $\left(1 - \frac{n_{vt}-1}{n_{vt}}\alpha_{vt}\right)$ is that (i) the measurement error in fertilizer or effort is uncorrelated with any of the independent variables, and (ii) there is no measurement error in the price of fertilizer. In this case, one can use OLS to estimate the following regression equation:

$$\log\left(\frac{f_{it}}{e_{it}}\right) = \varphi_i + \phi_{vt} - \sigma \log\left(p_{it}\right) + \epsilon_{it}, \tag{1.14}$$

where φ_i are household fixed effects and ϕ_{vt} are village-month fixed effects, which estimate $\sigma \log\left(\kappa_i\right)$ and $-\sigma \log\left(1 - \frac{n_{vt}-1}{n_{vt}}\alpha_{vt}\right)$, respectively. The identification of κ_i relies on the assumption that risk sharing is not household-specific; otherwise, φ_i would be also capturing variation in risk sharing at the household level. Also notice that I need both cross-sectional and time variation in fertilizer prices, because otherwise all their variation would be captured by the fixed effects. In fact, I observe dispersion in fertilizer prices across households and time. Possible explanations are costly information (Aker (2010)) and price discrimination operated by local traders.

The estimated elasticity of substitution between effort and fertilizer, $\widehat{\sigma}$, is about 0.35, confirming the complementarity between effort and fertilizer suggested by the evidence presented in Section 1.3.2.

In order to back out the marginal disutilities of effort, I compute $\widehat{k}_i = \exp\left\{\widehat{\log\left(\kappa_i\right)}\right\}$, where $\widehat{\log\left(\kappa_i\right)}$ is obtained by diving the household fixed effects by $\widehat{\sigma}$. Figure 1.1 shows the histogram of $\widehat{\log\left(k_i\right)}$.[24] The median marginal disutility of effort is approximately 3.7. To get a sense of this number, assume that households have quadratic utility. Then, the increase in consumption that would exactly compensate the median household for an increase in one hour of work (i.e., the marginal rate of substitution of effort for consump-

[24] I only show the estimates up the 85th percentile to make the graph more readable.

tion) is pinned down by the following equation:

$$\frac{\mathrm{d}c_i\left(\alpha\right)}{\mathrm{d}e_i} = \frac{3.7}{\rho c_i\left(\alpha\right)}.$$

As median household consumption is about 100 rupees, the increase in consumption compensating the median household for an extra hour of work is $0.037\rho^{-1}$ rupees. According to the estimates provided by the Indian government (Indian Labour Bureau (2010)), in 2009, the daily wage rate for an adult male agricultural worker fell in the range of 50 to 120 2009 rupees, which roughly correspond to an hourly wage rate (assuming eight hours of work per day) of 0.5 to 1.2 1975 rupees. If the labor market were competitive, then the marginal rate of substitution of effort for consumption would be equal to the hourly wage rate. This would imply a coefficient of absolute risk aversion between 0.074 and 0.031. These numbers are comparable with the coefficients of absolute risk aversion for medium stakes estimated by Binswanger (1981).

I can back out the wedges between the social and the private marginal benefit of effort using the same procedure employed to obtain the marginal disutilities of effort. Clearly, n_{vt} and α_{vt} cannot be separately identified. Nevertheless, following the standard practice in the literature (Ligon et al. (2002), Laczó (2015), Bold & Broer (2016)), I can set village size equal to the number of households sampled by ICRISAT and back out a structural estimate of risk sharing at the village-month level, $\widehat{\alpha}_{vt}$, by computing

$$\widehat{\alpha}_{vt} = \left(1 - \widehat{\zeta}_{vt}\right)\frac{\widetilde{n}_{vt}}{\widetilde{n}_{vt} - 1},$$

where $\widetilde{n}_{vt}$ is the imputed number of households sampled by ICRISAT. The number of households observed for each village and month is rather small: on average, less than 40 observations are used to compute a village-month fixed effect. This implies that $\widehat{\zeta}_{vt}$ is likely to be imprecisely estimated. By construction, $\zeta_{vt} \in [0, 1]$; however, half of the estimates of ζ_{vt} fall out of this range, being bigger than one. These observations cannot be used to estimate α_{vt}, because they would imply a negative sharing coefficient. The histogram of $\widehat{\alpha}_{vt}$ I obtain after dropping the estimates of ζ_{vt} that are bigger than one is given in Figure 1.2. On average, $\widehat{\alpha}_{vt}$ is equal to 0.7, but the estimates are more concentrated on the right of the distribution, with the median being equal to 0.8. Overall, these numbers square fairly well with existing empirical evidence on risk sharing in village economies, as well with the estimates obtained using the contrast estimator on my data (see Appendix A.4).

Counterfactual. How do fertilizer use and effort supply change when risk sharing changes? Consider Equation (1.13). Once parameters σ, κ_i, and n_{vt} are pinned down, I can move the sharing coefficients, α_{vt}, to quantify the effect of risk sharing on fertilizer used per hours worked. I use the structural estimates obtained above to pin down σ and κ_i. As for n_{vt}, I set village size equal to the number of households sampled by ICRISAT. Formally, I compute

$$\widetilde{x}_{it}\left(\widetilde{\alpha}_{vt}\right) := \widetilde{\log\left(\frac{f_{it}}{e_{it}}\right)} = \widehat{\sigma}\widehat{\log\left(\kappa_i\right)} - \widehat{\sigma}\log\left(1 - \frac{\widetilde{n}_{vt} - 1}{\widetilde{n}_{vt}}\widetilde{\alpha}_{vt}\right) - \widehat{\sigma}\log\left(p_{it}\right),$$

where $\widetilde{n}_{vt}$ is the number of households sampled by ICRISAT, $\widetilde{\alpha}_{vt}$ is imputed, and $\widetilde{x}_{it}$ is the resulting choice of fertilizer use per hours of work (in logs). Figure 1.3 shows the kernel density estimate of fertilizer used per hours worked when setting $\widetilde{\alpha}_{vt} = 0$ (black) and $\widetilde{\alpha}_{vt} = 1$ (grey). The summary statistics of $\widetilde{x}_{it}(0)$ and $\widetilde{x}_{it}(1)$ are reported in Table A8. On average, when going from no sharing to full insurance, the growth rate of fertilizer over effort is $1.75 - 0.36 = 1.39$; i.e., fertilizer used per hours worked more than doubles. The intuition behind this result is that both effort supply and fertilizer use decrease when moving from autarky to full insurance; however, effort supply is more responsive to changes in risk sharing than fertilizer use, and hence drops more than what fertilizer use does. This simple calculation quantifies the importance of risk sharing in shaping households' effort supply and fertilizer use.

The counterfactual exercise presented above quantifies the effect of risk sharing on fertilizer use per hours worked. Next, I disentangle the effect of risk sharing on effort supply and fertilizer use (see Appendix A.6). Table A9 reports the summary statistics of the growth rates of effort supply and fertilizer use when going from no sharing to full insurance. The average growth rates of effort supply and fertilizer use are -3.43 and -2.03. Thus, on average, fertilizer use decreases by four times and effort supply decreases by more than six times.

For which households are production decisions more affected by risk sharing? I regress household size, income, and plot area on the growth rate of fertilizer used per hours worked when moving from no sharing to full insurance. I find that, on average, fertilizer used per hours worked is more responsive for bigger households, and for households with higher monthly income and bigger plots (see Table A10). This evidence suggests that the choices of bigger and richer households are more affected by changes in risk sharing. A possible explanation is that bigger and richer households rely more on village-level risk-sharing arrangements than smaller and poorer households. This hypothesis is consistent with the evidence presented in Jalan & Ravallion (1999), which shows that poorer households are less well insured, and that household size is positively correlated with insurance against idiosyncratic income shocks in rural China.

Optimal risk sharing. Next, I compute the optimal sharing rule, solving Equation (1.6). Appendix A.7 reports the algebraic steps to solve this equation, and shows that I need values for χ, ρ, and η^2. Notice that my empirical strategy does not allow to retrieve these parameters. Hence, I proceed as follows. I build a grid of possible values for χ and ρ. In principle, $\chi \in [0, 1]$; however, for computational reasons, I take $\chi \in (0.1, 0.9)$.[25] and, following the evidence presented in Binswanger (1981), $\rho \in [0.001, 0.500]$. I set $\eta = 0.75$, following Morten (2017)'s estimate. Figure 1.5 shows the optimal sharing rule as a function of χ and ρ. The rows represent different values of ρ, and the columns represent different values of χ. The colors in the box represent different values of the optimal sharing rule: the darker a point, the closer to autarky. A first intuition is that when households are more risk averse it is optimal to give them more insurance: for a given χ, optimal sharing increases when moving to the right. The relationship between optimal sharing and land share is more complex and can exhibit a non-monotonicity. To see this, notice that, when ρ is sufficiently close to 0.001, the optimal sharing rule first decreases and then increases in χ. This non-monotonicity happens because the responsiveness of effort to the effective price of effort depends non-monotonically on χ.[26]

The effect of a fertilizer subsidy. The Indian government subsidizes fertilizer by assigning a so-called retention price to fertilizer. This price is fixed; i.e., it is independent of the quantity of fertilizer bought and sold in the market. The government pays the difference between retention price and sale price as subsidy to fertilizer manufactures for each unit sold. Hence, from the standpoint of the households, the government is exogenously changing the price of fertilizer. My model implies that a fertilizer subsidy will affect risk sharing and welfare (see Subsection 1.2.4). Figure 1.6 plots the optimal risk sharing rule (on the y-axis) against $s \in (0, 1]$ (on the x-axis), where I define s to be such that the price of fertilizer faced by household i in month t is sp_{it}.[27] Hence, one can see that higher fertilizer price leads to more risk sharing. For example, if the fertilizer subsidy is set to cut fertilizer price in half, my model predicts that risk sharing would

[25] To solve Equation (1.6), I make use of the bisection method, which is quicker but simpler than Newton's method. When χ gets close to 0 or 1, the bisection method does not perform well, as it is not able to compute any root.

[26] In particular, $\partial e_i^* / \partial p_e$ is first increasing and then decreasing in χ. Hence, when χ is sufficiently small, the cost of risk sharing (in terms of under-provision of effort) is increasing in χ, then becoming decreasing in χ.

[27] To draw this graph, I take $\rho = 0.01$ and $\chi = 0.6$. I choose these numbers because they imply an optimal sharing rule that is approximately equal to 0.8. Needless to say, there are other combinations of ρ and χ that give rise to similar optimal sharing rules. In fact, I plan to compute the effect of the subsidy on risk sharing for all possible values of ρ and χ.

decrease by 13%. The intuition is that, for the set of parameters estimated and calibrated, the slope of the effort supply function with respect to risk sharing becomes more negative when the price of fertilizer is lower. Thus, the subsidy increases the marginal cost of insurance, making it bigger than its marginal benefit. Because of the concavity of the welfare function around α^*, the planner decreases α to reestablish the equality between the marginal benefit and the marginal cost of risk sharing.

Another implication of my model is that if the countervailing effect of insurance on fertilizer use is not taken into account, a researcher would overestimate the elasticity of fertilizer to subsidy. The standard practice to estimate input demands in agricultural economics begins by specifying a translog production function, which can be conceived as a linear approximation to a CES production function. Then, Sheppard lemma is invoked to state that

$$\frac{\partial \pi_{it}}{\partial p_{it}} = f_{it}^* = \beta_f + \beta_{fy} \log\left(y_{it}\right) + \beta_{fe} \log\left(p_{it}^e\right) + \beta_{ff}\left(p_{it}\right),$$

where

$$p_{it}^e := \frac{\kappa_i}{\left(1 - \frac{n_{vt}-1}{n_{vt}}\alpha_{vt}\right)}.$$

Suppose that one does not take into account the effect of risk sharing on fertilizer use. Then, one would estimate

$$\frac{\partial f_{it}^*}{\partial p_{it}} = \beta_{ff}\frac{1}{p_{it}}.$$

On the other hand, my model suggests that

$$\frac{\partial f_{it}^*}{\partial p_{it}} = \beta_{fe}\frac{1}{p_{it}^e}\frac{\partial p_{it}^e}{\partial \alpha^*}\frac{\partial \alpha^*}{\partial p_{it}} + \beta_{ff}\frac{1}{p_{it}}.$$

Since effort and fertilizer are complements, $\beta_{fe} < 0$. Moreover, I have argued above that risk sharing is increasing in the price of fertilizer; i.e., $\partial\alpha^*/\partial p_{it} > 0$. Hence, noticing that $\beta_{ff} < 0$ and $\partial f_{it}^*/\partial p_{it} < 0$, one can see that

$$\frac{\partial f_{it}^*}{\partial p_{it}} < \beta_{ff}\frac{1}{p_{it}}.$$

Finally, Subsection 1.2.4 shows that a fertilizer subsidy increases welfare. I can compute the consumption-equivalent gain in welfare of a fertilizer subsidy; i.e., the percentage increase in aggregate consumption that would make the planner indifferent to switching back from the subsidized fertilizer price to the actual price. I find that the consumption-equivalent gain in welfare from cutting the price of fertilizer in half is 51%.

Validating the model. The reduced-form test for perfect risk sharing conducted in Subsection 1.3.2 indicates that the elasticity of consumption with respect to idiosyncratic income shocks is approximately 0.74 (see Table A2). If I take this elasticity as a proxy measure for α_{vt}, I can consider the following validation exercise: simulate fertilizer used per hours worked for $\widetilde{\alpha}_{vt} = 0.74$, and compare the simulated distribution with the actual distribution of fertilizer used per hours worked. If the income elasticity of consumption is a good proxy for α_{vt}, one should hope the simulated distribution to be able to match reasonably well the actual distribution. Figure 1.4 compares the two distributions. The black line is the density of simulated distribution and the grey line is the density of the actual distribution. The model matches the mean and the median of the actual distribution reasonably well, but the actual distribution is more dispersed than the simulated distribution, especially on the left tail. This suggests the presence of relevant sources of variation in fertilizer used per hours worked that are not present in the model.

A second validation exercise is the following. Calibrate the sharing rule to match a particular moment of the distribution of fertilizer per hours worked, and compare the resulting calibrated rule with estimates of α. If the model performs well, then the calibrated rule should be comparable to these estimates. First, I calibrate the sharing rule to match the mean of fertilizer per hours worked. To do so, I pick an α that minimizes the mean squared difference between $\log(f_{it}/e_{it})$ and $\log\widetilde{(f_{it}/e_{it})}$. The calibrated α is approximately 0.82, a number that squares very well with existing estimates of risk sharing. I also calibrate the sharing rule to match the median of fertilizer per hours worked. To do so, I choose an α that minimizes the mean absolute difference between $\log(f_{it}/e_{it})$ and $\log\widetilde{(f_{it}/e_{it})}$. In this other case, the calibrated α is approximately 0.75, which is extremely close to elasticity of consumption with respect to idiosyncratic income shocks estimated in Subsection 1.3.2.

A third validation exercise is as follows. Given the estimates of α obtained above, predict consumption using Equation (1.3). (Refer to this predicted consumption as simulated consumption.) Then, estimate Equation (1.7) using simulated consumption instead of actual consumption. If the model performs well, the estimated elasticity of simulated consumption with respect to idiosyncratic income shocks should be comparable to the estimates obtained in Subsection 1.3.2. Fist, I simulate consumption using the calibrated sharing rules calculated in the previous graph; then, I do so by using the estimated sharing rules backed out using $\widehat{\zeta}_{vt} := \left(1 - \frac{n_{vt}-1}{n_{vt}}\alpha_{vt}\right)$ and setting village size equal to the number of households interviewed by ICRISAT (I disregard the estimated sharing rules that do not lie in the $[0, 1]$ interval). Table A11 reports the OLS estimation of simulated consumption on actual income controlling for household and village-time fixed effects. Then, I simulate consumption using the estimated sharing rules computed in the previous paragraph. The elasticity of simulated consumption to idiosyncratic income shocks is positive and significant. When I use the α calibrated to match the mean of fertilizer use per hours worked the elasticity is 0.16, while when I use the α calibrated to match the median the elasticity is 0.21. These estimates are close to 0.26, the elasticity estimated using actual consumption in Subsection 1.3.2. However, my model seems to slightly underestimate the empirical loading on own income. Underestimation of income elasticity of consumption might be expected, as there probably are other relevant impediments to risk sharing other than private information frictions. When I use the the estimated sharing rules, elasticity is 0.29, which is slightly bigger than the actual elasticity. This slight overestimation might be due to the fact that the number of observations is significantly smaller.

1.4 Conclusions

While rural households in low-income countries face sizable random fluctuations in income, they often lack access to formal insurance. Despite this shortfall, these households manage to smooth their consumption, albeit imperfectly, by relying on informal insurance arrangements. Given their pervasiveness, these arrangements might well have an impact on technology adoption and agricultural input use. Studies on risk sharing abound, but few of them try to relate risk sharing to agricultural production decisions. In this chapter, I analyze the effect of informal insurance arrangements on fertilizer use when there are private information frictions in production decisions.

The chapter makes use of the following two insights. First, risk sharing can have a discouraging effect on households' incentives to exert effort. Second, effort and fertilizer are complementary inputs. The chapter outlines a static model of linear risk sharing which combines these two insights, and demonstrates theoretically that better insured households decrease their effort provision and fertilizer use. The model is useful in generating testable implications for the sign of the correlation between insurance and fertilizer use. One can easily test these implications with a many data-sets, as testing them only requires panel data on consumption, income, and agricultural production.

I test the model using the last ICRISAT panel from rural India. First, I implement a reduced-form strategy that confirms the main predictions of the model by showing that (i) insurance and effort supply are negatively correlated, (ii) effort supply and fertilizer use are positively correlated (which I interpret to be suggestive evidence of the existence of a complementarity between effort and fertilizer), and (iii) more insurance is associated to lower fertilizer use. Then, I structurally estimate the model. I obtain estimates for the elasticity of substitution between effort and fertilizer, the household-specific marginal disutility of effort, and the village- and month-specific constrained-efficient sharing rule. I use these estimates to quantify the effect of risk sharing on fertilizer use and effort supply. I find that when moving from no sharing to full insurance, average fertilizer use drops by four times and average effort supply decreases by six times. I also analyze the effect of a fertilizer subsidy on risk sharing and welfare. My model predicts that a 50% reduction in the observed price of fertilizer would generate a 13% drop in risk sharing and a 51% consumption-equivalent gain in welfare.

Other factors can play a relevant role in explaining farmers' effort supply and fertilizer use. Insurance itself may affect fertilizer use through channels different from the one I propose here. For example, Miller & Paulson (2007) argues that better insured households shift their portfolios toward riskier investments, and hence should be more willing to increase their use of riskier technologies. The fact that fertilizer might increase output volatility opens a new channel through which insurance might affect fertilizer use. While my model strips away from this channel, if the marginal impact of fertilizer on output volatility is not too big, then the results here presented remain valid. Other examples include Dercon & Christiaensen (2011), which finds that credit constraints and the possibly of low consumption when harvest is bad discourage fertilizer use; Beaman et al. (2013), which argues that a discrete change in the price or availability of fertilizer induces households to change other complementary inputs, making it difficult for them to isolate the return of fertilizer; Bold et al. (2017) which shows that low quality of fertilizer (possibly due to adulteration, poor storage, or inappropriate handling) significantly decreases the return of fertilizer, thus leading to low take up rates; and Duflo et al. (2011), which argues that present biases can induce farmers to procrastinate the decision to buy fertilizer, leading some of them to fail to invest in fertilizer altogether. These views and the one I propose here complement each other and together deepen our knowledge of the factors that lead farmers to under-utilize fertilizer.

While this chapter has focused on fertilizer use, many other production decisions can be affected by risk sharing. For example, in a very promising project, Mazur (2018) studies the joint determination of risk sharing and commonly owned agricultural equipment when there are limited commitment frictions. This project and this chapter are part of a very recent attempt to analyze the interaction between informal insurance and different household decisions in village economies (Advani (2017) and Morten (2017)). This recent surge of interest in the impact of risk-sharing arrangements on different facets of the village economy is motivated by the recognition that the design of these institutions can interact with households' decisions through different frictions at place. A fruitful avenue for future research may be to not only take into account the interaction of risk sharing and agricultural input use, but to simultaneously consider the presence of other frictions and production choices. For example, what if, as Mazur (2018) suggests, under limited commitment villages that invest more in irrigation technologies achieve higher levels of risk sharing? This chapter relies on a private information friction and argues that farmers that do more risk sharing work less hours and use less fertilizer. The coexistence of limited commitment and private information frictions might imply that risk sharing gives rise to a trade off between irrigation provision and fertilizer use. This is interesting because the agronomic literature suggests that fertilizer and irrigation are complements (Jaga & Patel (2012)); however, a combination of limited commitment and private information frictions might act as a force against this complementarity. If this was the case, risk sharing would induce farmers to use a sub-optimal combination of fertilizer and irrigation, which has consequences for agricultural yields and welfare.

Chapter 2

Propensity to Trust and Network Formation

2.1 Introduction

Many social interactions are influenced by how individuals sort into networks and groups (Demange & Wooders (2005)). From a policy perspective, the endogenous patterns of social interactions can have a profound effect on the outcomes of interventions that manipulate peer groups (Carrell et al. (2013)). There are many factors that play a significant role in network formation (Marmaros & Sacerdote (2006), Mayer & Puller (2008), and Wimmer & Lewis (2010)). This chapter analyzes how individuals' trust affects the probability that they form relationships.

Research in behavioral economics and neuroeconomics shows that trusting behavior in the laboratory might be related to factors that contribute to prosocial behavior in mammals. Specifically, Fehr et al. (2005) and Kosfeld et al. (2005) show that administration of oxytocin (a neuropeptide playing an important role in social attachment and affiliation in mammals) causes a substantial increase in trust among human subjects in a standard trust experiment. Moreover, research in sociology and social psychology argues that trust encourages individuals to approach strangers to form relationships (Yamagishi et al. (1998) and Igarashi et al. (2008)). It is thus reasonable to hypothesize that in a group of strangers who can socialize with each other, more trusting individuals form more relationships, other things being equal. In this chapter, we test this hypothesis.

The social capital literature has often put into relation trust and networks. In economics and sociology, trust has often been bundled into the very definition of social capital;[1] other times, measures of trust have been used as proxies for social capital (Kawachi

[1] Inglehart (1997) defines social capital as "a culture of trust and tolerance, in which extensive networks of voluntary associations emerge;" Putnam (1995) defines it as "features of social organization such as

(2010)). While it has been argued that both trust and networks play a role in determining social capital (Jackson (2018)), our aim is to shed light on the interconnection between trust and network formation, as in Carpenter et al. (2004).[2] On one hand, trust could help people to approach strangers to form relationships. On the other hand, having more relationships could push people to trust more.[3] Research in economics, sociology, and social psychology has demonstrated the existence of correlations between measures of trust and different network characteristics (Carpenter et al. (2004), Igarashi et al. (2008), and Carpiano & Fitterer (2014)). These correlations cannot be interpreted casually, as they are blurred by reverse causality: do people end up trusting more because they happened to form more relationships or do they form more relationships because they trusted more to begin with? In this chapter, we attempt to disentangle the effect of trust on network formation[4] by using an empirical strategy that is immune to reverse causality by design: we measure people's trust *before* they have a chance to socialize; then, we retrieve some of the networks they form and relate the trust measures collected to the networks retrieved. In order to deal with other sources of bias (i.e., measurement error and omitted variables), we make use of wide array of measures of trust, and control for many characteristics that might play a role in the formation of relationships and may be correlated to trust.

We collected data from a cohort of first-year economics undergraduate students at an elite university in Bogotá, Colombia. We put together the data in two stages. In the first stage, we measured trust for 72 members of the cohort before they had significant chances to get to know each other and socialize. Specifically, we measured trust on the "university welcome day," which is the first day in which students formally attend the university campus.[5] We measured trust in three ways: (1) a standard trust experiment (as in Berg et al. (1995)), (2) two survey questions adapted from the General Social Survey (GSS), and (3) three survey questions we designed to measure *particularized* trust; i.e., trust towards known people.[6] Moreover, we designed three survey questions to measure trustworthiness

networks, norms, and social trust that facilitate coordination and cooperation for mutual benefit;" and Woolcock (1998) defines it as "the information, trust, and norms of reciprocity inhering in one's social networks."

[2]According to Carpenter et al. (2004), the term 'social capital' is poorly defined because to one set of researchers it means the propensities of individuals to trust others, while to others it means the networks among individuals. The authors refer to the first definition as behavioral social capital and to the second as associational social capital.

[3]Glanville & Paxton (2007) provides evidence suggesting that people develop trust based on their past social experiences.

[4]Referring to Carpenter et al. (2004)'s classification, our aim is to understand how behavioral social capital impacts on associational social capital.

[5]In order to control for the fact that some of our subjects might know each other from before, we asked each subject to name each of the other subjects that he or she already knew.

[6]The social psychology literature refers to trust towards strangers as generalized trust, and generally contrasts it to particularized trust. In this chapter, we refer to trust towards strangers simply as trust, as it is customary in the economics literature.

towards friends and neighbors. During this stage, we also collected information on several demographic characteristics for the students using survey questions. After four months (i.e., at the end of the academic semester), we administered another survey to elicit five social networks among the students representing different relationships (greeting each other, friendship, studying together, having lunch together, and confiding in), and to get more demographic characteristics for the students.

Our data collection strategy was meant to reach three goals. First, we needed to measure the subjects' trust before they started socializing, and to be able to control for whether some of subjects knew each other from before socialization began. Second, we wanted the subjects to have substantial chances to socialize over an extended period of time.[7] Finally, we aimed for a group of people for whom trust could be measured accurately, and for whom information could be obtained on many other characteristics (including their social networks). Credibly estimating the causal impact of trust on network formation requires to minimize the impact of reverse causality, measurement error, and omitted variable bias. Our empirical strategy is virtually immune to reverse causality because we measured the subjects' trust before they had significant chances to socialize. In order to lessen the burden of measurement error and omitted variable bias, we needed to measure trust as accurately as possible, and to get information on characteristics that can affect network formation and might be correlated with trust. Obtaining accurate measurements of trust demands high levels of participation by the subjects. Our design allowed us to measure trust by means of both a standard trust experiment and survey questions. Moreover, we managed to ask several survey questions and combine the answers with administrative data coming from the university to gather information on many characteristics identified as relevant by the empirical literature on network formationSee e.g. Kandel (1978) and Marmaros & Sacerdote (2006). We discuss this literature more in detail below. and that we suspect to be possibly correlated with trust.

We estimate logistic regression models to identify how trust affects link formation probability in the networks we elicited. We use the quadratic assignment procedure (QAP) to account for possible correlation between unobservables affecting link formation.[8] We control for a large set of observable characteristics obtained through our survey and the administrative data from the university, including several individual and dyadic traits that may play an important role in network formation and might correlate with

[7] We retrieved the social networks at the end of the academic semester (i.e., four months after measuring trust) in the hope that stable, self-evident relationships would have enough time to form.

[8] See Good (2013) for an introduction to permutation tests, and Hubert & Schultz (1976), Krackardt (1987), and Krackhardt (1988) for the use of QAP. See Giné et al. (2010), Santos & Barrett (2010), Santos & Barrett (2011), and Giné & Mansuri (2018) for applications of QAP in economics.

trust. The main insight that comes out of our analysis is that trust poorly explains relationship formation for the networks we retrieve. Besides being generally insignificant, the effect of trust on link formation appears small with respect to other observable characteristics, such as homophily in socio-economic background. Specifically, the effect of trust is 3 to 4 times smaller than the effect of similarity in socio-economic status. Finally, with respect a model of network formation that does not control for trust, the model's explanatory power barely improves once we include trust. Our analysis delivers the result that trust does not play a relevant role in explaining our subjects' relationships after a socialization period of four months.

To summarize, our study serves as one of the first to identify the impact of trust on the formation of social relationships. Survey techniques (to elicit the students' attitudinal trust measures, social networks, and demographic characteristics) and controlled experimentation (to obtain behavioral trust measures), combined with an empirical strategy that is immune to reverse causality by design, allow for a credible estimation of the causal impact of trust on network formation, as they help reducing the impact of measurement error and omitted variable bias. More generally, our empirical strategy proves useful in estimating the causal impact of characteristics that may be endogenous to (or confounded by strategic interaction on) social networks on link formation.

Related Literature

Social scientists have long been interested in the role of trust on economic development. In one of the first systematic studies of trust, Edward (1958) argues that a lack of trust towards outsiders is responsible for economic backwardness in villages in Southern Italy. Arrow (1972) pushes this line of thought even further by stating that "much of the economic backwardness in the world can be explained by the lack of mutual confidence." Since these early contributions, a growing literature in economics and political science has conceptualized trust as one of the main determinants of economic development in market economies (Putnam et al. (1993) and Fukuyama (1995)). In more recent years, empirical research on the role of trust on macroeconomic outcomes has surged: survey measures of trust at the country level have been related to national economic outcomes, with the evidence suggesting that trust may be an important determinant of output growth, the inflation rate, and trade volume (Porta et al. (1997), Knack & Keefer (1997), Zak & Knack (2001), Guiso et al. (2009), and Falk et al. (2018)). The latest empirical research has begun to shed light on the specific channels by which trust matters for economic outcomes, by showing how trust affects financial development, the organization of firms, and labor market efficiency (see Algan & Cahuc (2013) for a review). We contribute to this literature by empirically investigating a new channel through which trust might affect economic outcomes; i.e., the effect of trust on network formation.

The literature on trust is also far-reaching in behavioral economics, sociology, and social psychology. Trust in strangers is widespread (Evans & Krueger (2009)), and it has been recognized to affect how people behave in strategic interactions in the laboratory (e.g., Bowles & Gintis (2002), Glaeser et al. (2002), Henrich et al. (2004), and Fehr (2009)). It has also been suggested that trust might be related to the formation of relationships and prosocial approach behavior (Yamagishi et al. (1998), Igarashi et al. (2008), Fehr et al. (2005), and Kosfeld et al. (2005)). If related to network formation (and thus to the place that we ultimately occupy in social networks), trust might not only affect *how* we play with a given set of players, but also *with whom* we play. We build on this idea and offer an empirical strategy to study the causal impact of trust on network formation. Our strategy is immune to reverse causality by design. In order to deal with measurement error in trust, we measure trust by using both controlled experimentation, mostly preferred by behavioral economists, and survey questions, most generally preferred by psychologists. Finally, to deal with omitted variables, we control for a large set of observable characteristics that are likely to play a role in network formation and may be correlated with trust, using a combination of survey and administrative data.

Economists have become increasingly interested in understanding how networks form (Demange & Wooders (2005)). While most work in this line of research has been theoretical (Jackson (2005)), economists have recently focused more attention on observational and experimental work on network formation (A. Chandrasekhar (2016)). From an empirical point of view, network formation rises questions such as how and why relationships form, why some people have fewer relationships than others, and whether people with similar traits are more likely to form relationships. Among the first contributions in this literature, Kandel (1978) analyzes a panel of adolescents and documents a strong connection between the presence of social ties and similarity in educational aspiration, political orientation, marijuana use, and engagement in delinquent activities. Marmaros & Sacerdote (2006) show that geographical and racial proximity are key determinants of friendships in a sample of students of Dartmouth College. Similarly, Mayer & Puller (2008) use Facebook data on Texas A&M college students and document that proximity in major, dorm, and race are significant proxies for friendship formation. A regularity that emerges from many empirical studies on network formation is homophily: the tendency of individuals to form connections with those like themselves (McPherson et al. (2001)). Consistent with this finding, we document that homophily along socio-economic background and gender is an important determinant of link formation. We contribute to the empirical literature on network formation by uncovering the role of trust in the formation of social relationships among college students.

This work also relates to a recent attempt to connect laboratory experiments with social networks. Most of the papers in this line of research aim to understand how a given network structure influences play in an experimental setting. For example, Leider et al. (2009) elicit a friendship network of college students and show that dictators in a dictator game donate more to friends (recipients with social distance equal to one). Similar papers are Branas-Garza et al. (2010), which finds that betweenness centrality and reciprocal degree are positively correlated with giving in a dictator game; Goeree et al. (2010), which finds evidence of an inverse distance law of giving in a dictator game;[9] and Kovářík et al. (2012), which finds a positive relationship between subjects' inequality aversion and several centrality measures. To the best of our knowledge, this chapter is the first aiming to isolate the impact of behavior in an experimental setting on network formation.

Finally, this chapter is connected to a recent effort to link laboratory experiments with survey methods (Glaeser et al. (2000), Fehr et al. (2003), Falk et al. (2016), and Falk et al. (2018)). In particular, following Glaeser et al. (2000) and Falk et al. (2016), we measure trust by combining the standard trust experiment presented in Berg et al. (1995) with attitudinal survey questions about trust.

2.2 Design and Protocols

We collected data from first-year undergraduate students in economics at an elite university in Bogotá, Colombia. Our design consisted of two stages. The first stage was conducted on August 4, 2017 and its main goal was to measure students' trust. Crucially, this stage was carried out on the "university welcome day," which is the very first day in which students formally attend the university campus. The idea was to measure students' trust before they had significant opportunities to socialize. The second stage was conducted online between December 7, 2017 and January 5, 2018 and its main aim was to elicit social networks among the students, at the end of their first semester of classes. In what follows we describe the design of the two stages in detail.

First stage. The first stage was aimed at the 81 students comprising the economics class of the first semester of 2017, and it was implemented using the paper handouts labeled A, B, C and D, included in Appendix B.2.[10] The questionnaire was administered during a session that took place right after the students had lunch and lasted 90 minutes. 72 of the

[9] I.e., a person who is set to play as the dictator with someone who is further away in the network gives progressively less.

[10] The session was conducted in Spanish. Appendix B.2 includes the original handouts. See `https://sites.google.com/view/davide-pietrobon/research?authuser=0` for an English translation.

81 students summoned attended the session. Handout A contains a general description of the activity, along with the informed consent form required for participating in the session. Handout B is a detailed description of the trust game. Handout C is a form for recording the students' strategies in the trust experiment for each of the two roles (i.e., sender and receiver). Finally, handout D is a questionnaire with two questions (4(a) and 4(b)) aimed to measure the students' generalized trust, six questions (5(a)-5(e)) aimed to measure their particularized trust, and six questions (1-3 and 6-8) on individual characteristics.[11]

In the trust experiment, every participant was endowed with $e = 20,000$ Colombian Pesos (COP) (about $USD\$7$). In every anonymously created sender-receiver pair, each sender (he) had to decide how much money, s, to transfer to the receiver in a range from 0 to e in $\Delta := COP\$2,000$ increments. The receiver (she) would receive three times the money sent to her by the sender, $3s$, and, contingent on the amount received, r, had to decide how much money to send back to the receiver $f(r)$. In each case, she could send back any amount in a range from 0 to r in Δ increments. The monetary payoffs at the end of the game for a sender-receiver pair in which the sender uses strategy s and the receiver uses strategy $f(r)$ are $e - s + f(3s)$ to the sender and $e + 3s - f(3s)$ to the receiver.

The two roles in the trust experiment were described to all participants, and they were informed that each of them was required to report how they would behave both roles (sender and receiver), as we would then assign these roles randomly[12] and match them randomly to each other, to implement their reported strategies and determine monetary payoffs.[13]. Participants were instructed that all actual monetary payments would be made two weeks after the experiment was undertaken.[14] Handout B included instructions for the strategies available to the sender and the receiver, a description of the functions used to calculate the monetary payoffs, and a detailed example. The instructions and the example were read out aloud, and a question-and-answer session was conducted right afterwards. The students were then instructed to fill out handout C, which contained the actual strategies to be chosen. Specifying the strategy for the role of sender entailed

[11]These were sex, age, number of siblings, number of friends, number of people in the cohort that the person knew from before, and self-reported well-being.

[12]Half of the subjects were assigned to be senders and the other half to be receivers. In case of an odd number of participants (this was not the case) we planned to include an extra "artificial" player relying on a pair of strategies (one for each role) randomly chosen from the set of strategy pairs actually submitted by the subjects.

[13]We chose to have every participant assume both roles in order to record the behavior of as many senders as possible.

[14]For administrative reasons the payments were in fact only made 4 months later, at the end of the academic semester.

stating one among 11 (0-10) transfer options in Δ units. Specifying the strategy for the role of receiver entailed stating 11 *contingent* transfers, one for each of the 11 possible amounts received from the sender. In each case, the receiver could choose to send back to the sender an amount ranging from 0 to the entire amount received in Δ units.

The students also filled out a survey contained in handout D, with 8 questions aimed to measure generalized trust, particularized trust towards friends and neighbors, and particularized trustworthiness towards friends and neighbors. We report the questions below.

4- To what extent do you agree with the following statements (on a 1-5 scale, where 1 denotes total disagreement and 5 total agreement)

(a) One cannot trust strangers (Henceforth A2)

(b) When dealing with strangers it is important to be careful and not to readily trust them. (Henceforth A3)

5(a) To how many among your 10 closest friends have you lent money? (Henceforth B1)

5(b) How many among your 10 closest friends have lent money to you? (Henceforth B2)

5(c) To how many among your 10 closest friends have you lent your belongings (e.g., books, CDs, clothing, bicycle)? (Henceforth B3)

5(d) How many among your 10 closest friends have lent their belongings (e.g., books, CDs, clothing, bicycle) to you? (Henceforth B4)

5(e) How many among your 10 closest neighbors would you trust with your house keys? (Henceforth C1)

5(f) How many among your 10 closest neighbors would trust you with their house keys? (Henceforth C2)

Questions A2 and A3 measure generalized trust, questions B1 and B3 measure particularized trust towards friends, question C1 measures particularized trust towards neighbors, questions B2 and B4 measure particularized trustworthiness towards friends, and question C2 measures particularized trustworthiness towards neighbors. The survey also included five questions on demographic characteristics (sex, age, number of siblings, number of friends, number of people in the cohort knew from before) and one question on self-assessed well-being.

Second stage. The second stage of the data collection process was conducted four months later, and its goal was to elicit some of the networks of social relationships among the 81 students that participated in the first stage. Additionally, we asked the participants a number of questions on individual characteristics that may play role in network formation and might be correlated with trust. For this purpose, we invited 113 students to complete an incentivized survey.[15] The set of 113 students consisted of the 81 intended subjects of the first stage along with 32 other students who, due to their course schedules, were in frequent contact with the 81 intended subjects throughout the semester.[16] The elicitation of the student's social networks was done in two parts. First, the students were presented with the complete list of the 113 invited students (in random order), and were asked to identify the people who they greeted (henceforth, hello partners). Specifically, for each student in the list they were asked to tick a box if they would say hi to that student upon encountering him or her. Secondly, each student was presented with his or her complete list of hello partners and, for each of them, he or she was asked to check one or more of six boxes acknowledging the following relationships: (1) "I am friend of this person," (2) "I frequently study with this person," (3) "I frequently have lunch/coffee with this person," (4) "I share personal matters with this person," (5) "I was acquainted with this person before entering University," and (6) "None of the options (1)-(5) apply to my relationship with this person." Thanks to box (5) we are able to control for whether relationships formed before our intended socialization period, and so we end up with five possible relationships (greeting each other, friendship, studying together, having lunch together, and confiding in).

Besides containing questions meant to elicit networks, the survey included several questions on individual characteristics (physical and psychological) that we suspect to possibly play a role relationship formation and correlate with trust. These were number of siblings, number of friends enrolled in the same university met before starting university, number of friends enrolled in the same university met after starting university, number of friends not enrolled in the same university, average weekly hours spent socializing with friends enrolled in the same university, average weekly hours spent socializing with friends not enrolled in the same university, average weekly hours spent doing physical activities,

[15] We offered each respondent a COP\$20.000 (approximately USD\$7) voucher for a fast food restaurant on campus. Appendix B.2 includes the original survey, which was conducted in Qualtrics. See `https://sites.google.com/view/davide-pietrobon/research?authuser=0` for an English translation of the original survey.

[16] The intended subjects had to take 6 courses, out of which 5 were mandatory. The 32 remaining students were students from other semesters who also took the course "Pensando Problemas" (one of the mandatory courses) that same semester together with the 81 intended subjects. With rare exceptions, the 32 remaining students comprehended students who decided to declare themselves as students of the economics program later in their university careers, or students who failed and had to retake the "Pensando Problemas" course (these amount to about 10% of the students in each cohort). Although our focus was on the 81 students that made up the incoming class, we decided to include the other 32 students to have more observations on relationship formation. With very few exceptions, the 81 intended subjects had no prior acquaintance with the other students.

hobbies, hometown, age, eye color, hair color, height, weight, whether wearing glasses, whether wearing tattoos, whether wearing piercings, whether smoking, whether attending parties, place of living in Bogotá, and four personality questions. In these questions, the students were asked to rate on a scale from 1 to 5 how much they perceived themselves as realistic, introverted, inhibited, and shy. Finally, we asked the students to rate on a scale from 1 to 5 how much they agreed with the following statements: "I am very sociable," "I am satisfied with my social life," "making friends at university is easier than I thought," and "I am satisfied with the number of friends I have." In addition to the data collected with these survey questions, our empirical analysis uses administrative data on a large set of students' characteristics, such as the scores obtained at the high school national exit examination, their GPAs, their socio-economic background, and the amount of time they spent together in the same classrooms during the first semester. See Subsection 2.3.1 for a more detailed discussion.

2.2.1 Student Characteristics and Networks

Table B.1 contains summary statistics for the entire population.

The average height is 172.1 cm (ranging from 60 cm to 192 cm) and the average weight is 65.4 kg (ranging from 46 kg to 100 kg). Ages range from 16 to 26 years old (with an average of 18.4); the number of siblings ranges from 0 to 4 (average 1.3); 30% wear glasses, 4% has tattoos, 19% wear piercings, 32% are smokers, and 85% report going to parties. The questions on personality characteristics show answers ranging from 1 (corresponding to the left-most bubble in the response form) to 5 (corresponding to the right-most bubble in the response form) with means of 3.2 (realistic), 2.9 (introverted), 2.5 (inhibited), and 2.3 (shy), respectively. On average, subjects spend 10.4, 7.3, and 4.4 hours per week socializing with friends enrolled in the same university, socializing with friends not enrolled in the same university, and doing physical activity. The average number of friends enrolled in the same university met before starting university is 10.7 (ranging from 0 to 50), the average number of friends enrolled in the same university met after starting university is 9.9 (ranging from 0 to 40), and the average number of friends not enrolled in the same university is 16.6 (ranging from 0 to 100).

In the trust experiment, on average, the sender sent \$9811.32 COP out of his endowment of COP\$20,000 (standard deviation is COP\$5251.58). Survey questions on generalized trust exhibit means of 3.4 ("one cannot trust strangers" (A2)) and 3.89 ("when dealing with strangers it is important to be careful and not to readily trust them" (A3)); questions on particularized trust exhibit means of 5.53 ("to how many among your 10 closest friends have you lent money?" (B1)), 5.51 ("how many among your 10 closest friends have lent money to you?" (B2)), 6.62 ("to how many among your 10 closest friends have you lent your belongings?" (B3)), 5.98 ("how many among your 10 closest friends have lent their belongings to you?" (B4)); questions on trust towards neighbors exhibit means of 1.89 ("how many among your 10 closest neighbors would you trust with your house keys?" (C1)) and 2.1 ("how many among your 10 closest neighbors would trust you with their house keys?" (C2)). Figure 2.1 shows the correlations between the answers to questions A1, A2, A3, B1, B2, B3, B4, C1, and C2, where A1 is the amount of money sent in the trust experiment (i.e., behavioral trust). There is significant correlation between behavioral trust and generalized trust, as measured by questions A2 and A3, which are also significantly correlated. Particularized trust and trustworthiness towards friends, as measured by questions B1, B2, B3, and B4, are highly correlated, and so are particularized trust and trustworthiness towards neighbors, as measured by questions C1 and C2; however, trust and trustworthiness towards friends are not significantly correlated with trust and trustworthiness towards strangers.

Figures 2.2–2.4 display the five networks we retrieved (greeting each other, friendship, studying together, having lunch together, and confiding in) and the network of subjects who knew each other from before our intended socialization period. A thin (light grey) directed arrow is drawn from subject i to j when i names j.[17] Table B.2 displays some summary statics for the social networks.

2.3 Empirical Analysis

Our hypothesis is that subjects that exhibit higher levels of trust form more relationships, other things being equal. To test this hypothesis, we use logistic regression models to estimate the impact of trust on the probability of forming links in the networks we elicited. We control for individual characteristics, and allow for homophily along a variety of dimensions that may be relevant for network formation and might be correlated to trust. We use QAP to obtain correct p-values.

For each of the five relationships elicited, we can consider three versions of the network (G): the directed network (as reported in the survey), the union ($ij \in G$ if i nominates j or j nominates i), and the intersection ($ij \in G$ if i nominates j and j nominates i). Our baseline strategy focuses on the directed network; when we repeat the analysis for the union and the intersection the results are virtually unchanged.[18] Moreover, we can use (at least) six different measures of trust: the money sent in the trust experiment,

[17] For comparability, the nodes in all networks are displayed using a Fruchterman-Reingold type layout of the "greeting each other" network.

[18] Results are available upon request.

the answers to the two survey questions on generalized trust, and the answers to the three survey questions on particularized trust. Our baseline strategy uses the money sent in the experiment; when we repeat the analysis using the other measures of trust our conclusions are unaffected.[19]

The basic regression equation is as follows:

$$\log\left(\frac{\Pr\left(ij \in G \mid X_i, X_j, Z_{ij}\right)}{\Pr\left(ij \notin G \mid X_i, X_j, Z_{ij}\right)}\right) = \beta_0 + \beta_1 X_i + \beta_2 X_j + \beta_3 Z_{ij}. \tag{2.1}$$

where X_i and X_j are individual characteristics (the students' trust, sex, socio-economic background, and so on), and Z_{ij} are link characteristics, such as homophily[20] along physical characteristics (such as height and weight) and demographic characteristics (such as age, sex, place of living in Bogotá, and socio-economic background); whether they knew each other from before, time of exposure in class, and so on. Subsection 2.3.1 presents a careful description of all the controls introduced in the regressions and the way in which they are constructed.

Tables B.3-B.7 report the results of our baseline strategy using the five *directed* networks. The unit of analysis is the pair of individuals i, j. For each pair i, j, the variables called "Trust i" and "Trust j" are the amounts of money sent by individuals i and j in the trust experiment in their role of senders. These amounts are measured in $COP\$2.000$ units (from 0 to 10). In order to control for the possibility of homophily in trust, we include a variable called "Trust Δ," which is the absolute difference between the amounts sent (in $COP\$2.000$ units) by i and j. The coefficients associated to the controls are omitted for readability.[21] The tables show that trust has not a significant effect on the formation of the networks retrieved.

2.3.1 Controls

In order to reduce omitted variable bias, we control for a large set of individual and dyadic characteristics that may affect network formation and might be correlated with trust. As for individual characteristics, we control for sex, age, eye color, hair color, height, weight, whether wearing glasses, whether wearing tattoos, whether wearing piercings, number of siblings, number of friends enrolled in the university met before starting university, average weekly hours spent doing physical activity, whether attending parties, score obtained at the high school national exit examination, GPA, self-assessed personality type

[19]Results are available upon request.
[20]See Subsection 2.3.1.
[21]Appendix B.1 reports the tables including the coefficients associated to all the controls in the regressions.

(shy, inhibited, introverted, realistic), socio-economic background,[22] and place of living in Bogotá. As for dyadic characteristics, we control for whether the students reported knowing each other from before our intended socialization period, and for the amount of time they spent together in the same classrooms during the first semester, as measured by the number of university credits that the students share.[23]

The regression tables with the estimates of the coefficients associated to all the controls can be seen in Appendix B.1. For each control x and each pair of individuals i, j, the variables called "x i" and "x j" indicate the values of the control for individuals i and j. The variable called "x Δ" indicates the absolute difference between the values of the control for i and j if the control is a cardinal variable,[24] and a dummy indicating whether i and j share the same characteristic if the control is an ordinal variable. We introduce this variable to take into account the possibility of homophily in all of the individual characteristics mentioned above.

As expected, the time of exposure and the fact that the individuals knew each other from before the intended socialization period have a significant and sizable impact on network formation. The (average) marginal effect of knowing each other from before can be as high as 67% for the network of greetings; its effect is smaller for the other networks, ranging from 8% to 26%, but still significant at the 1% level. The marginal effect of time of exposure is smaller relative of the effect of knowing each other from before but still strongly significant. Not surprisingly, having an introvert personality has a significant negative effect on the formation of relationships. Finally, homophily is socio-economic background is a significant determinant of network formation, and its effect can go so far as being one order of magnitude bigger than the effect of trust.

2.3.2 Robustness Checks

We conduct a number of robustness checks in order to address potential concerns with our main specification.

A major concern with using logistic regression models such as Equation (2.1) is that they do not take into account network effects. For example, the presence of friends in common might ease friendship formation, or individuals may prefer to set up relationships with people who are better-connected or more central in the network. If such effects exist, the unit of analysis for estimating a network formation model cannot be the pair

[22] Socio-economic background is proxied by an administrative classification referred to as "estratificación socioeconómica" (socio-economic stratification), which classifies residential real estates into six categories, ranging from 1 (corresponding to the poorest socio-economic stratus) to 6 (corresponding to the richest one).

[23] Class sections are assigned randomly to first semester students.

[24] See e.g. Mouw (2006).

of individuals but should instead include the entire network.[25] The standard practice to deal with network effects is to use exponential random graph models (ERGMs). These models are extremely useful because they can incorporate any form of interdependency in relationships, but are far from problematic as the maximum likelihood estimator of their parameters can be inconsistent or computationally unfeasible (A. G. Chandrasekhar & Jackson (2014)). Nevertheless, we estimate ERGMs to control for several network effects, and the results we get stay in line with the findings of our main specification (see Appendix B.3).

Another concern is that the behavioral measure of trust given by the amount of money sent in the trust experiment may be subject to measurement error or might not capture the type of trust that matters in establishing relationships. To address this issues, we modify our baseline strategy by considering alternative measures of trust, using the answers to the surveys questions about both generalized and particularized trust. We also build a trust index by performing a principal component analysis of the amount sent in the experiment (A1), the answers to the survey questions about generalized trust (A2, A3), and the answers to the questions about particularized trust (B1, B3, and C1), and repeat our baseline strategy using this index as the relevant measure of trust. In all of these cases, we find that the effect of trust on the relation formation is small and typically not significant at the 5% level, confirming the results obtained with our baseline specification.

2.3.3 Discussion

All in all, we believe that there are two ways in which our results can be rationalized besides concluding that trust does not matter for network formation.

First, our results may suffer from omitted variable bias, even if we control for a very large set of individual and dyadic characteristics. In particular, late adolescents may value traits like dominance, charisma, or nerve when forming relationships with their peers. These traits could in turn be negatively correlated with trust. This hypothesis stands in line with the literature on social status among adolescents. Parkhurst & Hopmeyer (1998) distinguishes between sociometric popularity, representing the in-degree of an individual in a network,[26] and perceived popularity, representing an individual's reputation for being popular. Later studies (Hawke & Rieger (2013) and Franken et al. (2017)) distinguish also between the perception of being popular (popularity) and the perception of being well-liked (likability). The in-degree of adolescents is consistently found to be positively correlated with both likability and popularity. However,

[25]See Robins et al. (2007) and references therein.
[26]In a directed network, the in-degree of an individual i is the number of links to i.

while likability is mostly related to pro-social behavior and traits, popularity is primarily associated with social dominance (Parkhurst & Hopmeyer (1998)) and correlated with physical aggression, relational aggression, and anti-social behavior (Cillessen & Mayeux (2004), LaFontana & Cillessen (2002), and Hawke & Rieger (2013)). Popular adolescents are also described as manipulative, Machiavellian (Cillessen & Mayeux (2004)) and hard to push around (Parkhurst & Hopmeyer (1998)). We hypothesize that while trust may be positively related to likability, it may have a stronger negative relationship with popularity.

Another possibility is that trust may be a relevant determinant of network formation in the long run but not in the short run. In particular, it could be the case that individuals initially form relationships based on similarity along observable characteristics, such as time of exposure and socio-economic background. Nevertheless, as time goes by, more trusting individuals might be more successful in maintaining a higher number of their initial relationships because they exhibit more pro-sociality. If this were the case, trust might have no significant effect on relationship formation for the networks we retrieve simply because we did not wait long enough to retrieve the networks.

2.4 Conclusions

We collected survey and experimental data on trust, as well as survey data on individual characteristics, from a cohort of first-semester undergraduate students in economics at an elite university in Bogotá, Colombia. These data were collected on the "university welcome day," before the students had significant opportunities to socialize. After four months, we collected survey data on several social networks among the students. We estimate network formation models for each of the networks elicited to identify how trust affects link formation. We control for a large set of observable characteristics obtained through survey and administrative data.

We find that trust poorly explains the formation of the networks we retrieve. In particular, the effect of homophily in socio-economic background can go so far as being one order of magnitude bigger than the effect of trust. This result suggests that trust towards strangers is far less important than other observable characteristics, such as time of exposure and socio-economic background, in establishing the patterns of social interactions among our subjects.

Very broadly, this chapter contributes to the burgeoning literature on the determinants of network formation. To the best of our knowledge, this work is the first to study how certain traits revealed by behavior in a controlled experimental setting affect how subjects form relationships in the field. More specifically, this is the first work trying to isolate the causal impact of trust on network formation by (1) implementing an empirical strategy that is immune to reverse causality by design; (2) controlling for a large set of observable characteristics, which are likely to play a role in network formation and might be correlated to trust, to reduce omitted variable bias concerns; and (3) combining survey and experimental measurements of trust to lessen the impact of measurement error.

The chapter also contributes to the literature on social capital by providing a first study of how behavioral social capital (in the form of trust) impacts on associational social capital (in the form of social networks). From this point of view, the chapter suggests that behavioral social capital is not an important determinant of associational social capital. While this result might seem puzzling, there are at least two explanations that can rationalize this fact. First, it might be the case that the positive correlations between measures of behavioral and associational social capital found in the literature (see, e.g., Carpenter et al. (2004)) hide a causal relationship that goes from associational to behavioral social capital (as suggested by Finseraas et al. (2019)); i.e., people tend to trust more because they have more friends. Second, it might be the case that trust does not matter in the short run but matters in the long run. For example, it might be the case that people initially sort themselves into networks and groups based on observable characteristics (e.g., homophily along socio-economic backgrounds); however, after a sufficient amount of time, people that behave more trustingly manage to maintain more of the initial relationships relative to people that trust less. If this were the case, trust would have no impact on the formation of the networks we retrieve four months after the measurement of trust, but would have a significant effect on future networks among the subjects.

Chapter 3

Better Not to Know: Uncertainty and Team Formation

3.1 Introduction

Consider a university department composed of two researchers named Alice and Bob, who are each starting to work on a new project. A researcher's performance is given by the quality of his research project. Before they start working on their projects, Alice and Bob can agree to give feedback on each other's projects (i.e., collaborate). If Alice's feedback helps Bob to improve his project then it increases the quality of Bob's project. In this case, Alice is said to be a peach. Otherwise, Alice's feedback is at best a loss of time for Bob, and at worst it will derail his project to a dead end, so it decreases the quality of Bob's project. In this case, Alice is said to be a lemon. Alice turns out to be a peach for Bob as long as her ability is higher than the difficulty of Bob's project (in which case she's able to give constructive feedback). The same goes for Bob's feedback on Alice's project. Alice and Bob are risk neutral and only care about their own performance. They have worked in the department for a long time so that they have come to know each other's ability.

To begin with, imagine that Alice and Bob know the difficulty of each other's project before they start working on them. Then, at the moment in which they may agree to collaborate, there is perfect information about whether a researcher is a peach or a lemon. The researchers would collaborate only in the case in which both are peaches because none of them wants to receive feedback from a lemon. Next, imagine that there is imperfect and symmetric information about the difficulty of their projects before they start working on them.[1] In this case there is public uncertainty about whether a

[1] Imperfect information naturally arises if the researches cannot perfectly forecast how difficult a project will turn out to be before they start working on them. Assuming that information is symmetric

researcher is a peach or a lemon. In this case, the researchers could find collaborating ex ante profitable. Ex post, the researchers could be collaborating even in the case in which one of them is a lemon. When looking at aggregate performance (that is, the sum of the researchers' performances), uncertainty about the difficulty of their projects could be desirable. This would happen if the increase in performance that a researcher enjoys when collaborating with a peach is higher than the loss in performance that he suffers when receiving feedback from a lemon. In this case, public uncertainty can make a peach ex ante willing to collaborate with a lemon, so that ex post aggregate performance would be higher than if there was perfect information (in which the researchers would not collaborate).

In general, when agents may agree to cooperate to confer externalities on each other, public uncertainty about the private returns to cooperation might be a tool to incentivize them to cooperate. If there are states of the world in which cooperation is socially desirable yet incentive incompatible, then public uncertainty about the private returns to cooperation may be socially beneficial.

I outline a model to test this theory. A risk neutral principal hires two-risk neutral agents to each work on a task. The agents can agree to help one another to fulfill their tasks (team up). They hold symmetric and imperfect information about whether an agent has the ability to help the other in performing his task (whether an agent is a good or bad teammate). When they team up, an agent receives a positive productive externality if his teammate is good and a negative externality if he is bad. When they do not team up, each agent's output is normalized to 0. The principal wants to maximize total output. Collaboration is not verifiable so that each agent is paid for the level of output obtained in his task. The principal can draw a noisy public signal of whether the agents have the abilities to help one another to fulfill their tasks before the agents may agree to team up. The noise of the signal is the probability that it gives a Type 1 error.[2] What happens to the agents' incentives to team up when the principal draws a signal? What happens to expected total output? Can drawing a signal decrease expected total output?

When noise is sufficiently high, drawing a signal does not affect the agents' incentives to team up. In this case, the signal is irrelevant. If noise is sufficiently low, when drawing a signal the agents team up as long as the signal tells that they are both good. I define the gain of drawing a signal as the difference in expected aggregate output when the principal draws a signal and when he does not. Drawing a relevant signal can decrease

means that there are no asymmetries in the researchers' beliefs about the difficulty of each other's projects.

[2]I.e., the noise of signal is the probability that it tells that an agent is good (bad) given that he is bad (good).

expected aggregate output, and the relationship between gain and noise can have at most one discontinuity and be non-monotonic. I characterize how gain changes as a function of noise for the case in which the prior probability that an agent is good is sufficiently high.

When there is public uncertainty about whether an agent is good or bad, there is public uncertainty about the externalities unleashed by cooperation and, as a result, about the private returns to cooperation. In this chapter, I analyze the effects of public uncertainty about the utility of teaming up on incentives and expected social welfare.

From a positive point of view, there are many situations where agents may agree to cooperate under public uncertainty about the private returns to cooperation. Imagine a client commissioning a market analysis report to a consulting firm. Writing the report involves analyzing market data from two countries, Italy and Spain. Two consultants work on the report, one analyzing data from Italy and the other analyzing data from Spain. The consultants can agree to help each other to fulfill their tasks. When they cooperate, a consultant increases his output if his colleague has the ability to help him fulfilling his task and decreases his output otherwise. Suppose that the agents' abilities are common knowledge but that the specific features of the tasks to be fulfilled are unknown so that there is public uncertainty about the difficulty of their tasks. At the moment in which the consultants may agree to help each other, there is public uncertainty about whether a consultant is a good or bad teammate. If each consultant is paid for the output he produces (and he only cares about his wage) then there is public uncertainty about the private returns to teaming up. The same reasoning applies when two (selfish) political parties may agree to form a coalition while not knowing whether a party's political platform will be supported by the electorate, when two firms may form an R&D partnership to develop a project while not knowing the amount of firm-specific human capital and infrastructure needed to develop the project, or when two economists may agree to give feedback on each other's work while not knowing which competencies are needed to give valuable suggestions.

When agents team up, they confer externalities on each other. Externalities are, by definition, non-contractible: agents cannot use side-payments to reallocate the surplus of cooperation. This can create tension between the agents' incentives and welfare. To see this, suppose that (i) agent 1 is a good teammate, while agent 2 is a bad teammate; and (ii) the size of the positive externality that agent 1 confers on agent 2 if they team up is 6, while the size of the negative externality that agent 2 confers on agent 1 is 3. If the agents team up then agent 1's payoff is -3, while if they do not team up then his payoff is 0. Since agent 1 does not agree to team up, there is no cooperation, and welfare is 0. However, the welfare of a team is $-3 + 6 = 3$: without transfers, the stable outcome is not optimal. What happens when utility is transferable? If agent 2 transfers at least 3 utils to agent 1 then agent 1 agrees to team up. Notice that agent 2 would agree to transfer at most 6 utils to agent 1. Thus, for each transfer $x \in [3, 6]$ that agent 2 makes to agent 1, the agents would agree to team up and social welfare would be $(-3 + x) + (6 - x) = 3$. Uncertainty about the utility of teaming up is interesting from a normative point of view because it imposes transfers of ex ante utility between the agents. To see this, suppose that the probability that an agent is good is 0.5. The expected utility of teaming up is $0.5 \times 6 + 0.5 \times (-3) = 1.5$. Under uncertainty about the utility of teaming up, agent 2 has to "transfer" 4.5 expected utils to agent 1, relative to the perfect information case. Ex ante, the agents team up, which is what a utilitarian social planner would like them to do. This shows that there are cases where the transfers of ex ante utility imposed by uncertainty bring about an ex ante alignment of the agents' incentives with social goals. In other words, uncertainty about the utility of teaming up can make ex ante incentive compatible a socially desirable team structure that is ex post incentive incompatible. In these cases, uncertainty about the utility of teaming up plays a similar role to randomization in resource allocation.[3]

Related Literature

This work is closely related to the literature on the value of public information. The earliest contribution in this literature is Hirshleifer (1971). He shows that in a pure exchange economy if agents are risk averse then drawing a public signal of the state of the world can only decrease expected social welfare. The reason is that if agents trade in complete markets for contingent claims before the state realizes then they share risks. Drawing a public signal reduces the agents' trading opportunities and hence risk sharing. More recently, Schlee (2001) generalizes Hirshleifer (1971)'s result by showing that in a pure exchange economy, if agents are risk averse then drawing a public signal reduces expected social welfare under one of the following assumptions: there is no aggregate risk, some agents are risk neutral, or there is a representative agent. In the risk sharing literature, this is known as the "Hirshleifer effect." I contribute to this literature by showing that drawing a public signal of the state of the world can decrease expected social welfare even if the agents are risk neutral. The reason why this happens is that in my model inefficiencies do not necessarily arise from reduced risk sharing but from the fact that, when one agent is good and the other is bad, a team can be socially desirable but incentive incompatible. Public uncertainty about the utility of teaming up can relax the good agent's incentive compatibility constraint up to the point where he agrees to

[3]See Budish et al. (2013) for an introduction to the theory and applications of random allocation mechanisms.

cooperate.

Morris & Shin (2002) outline a model to assess the value of a public signal of the state of the world. In their view, a public signal is a "double-edged instrument," conveying information on the underlying state on one hand and serving as a coordination mechanism on the other. They show that if agents have no private information then drawing a public signal always increases expected social welfare, and a more precise signal is always better than a less precise one. I show that drawing a public signal can decrease expected social welfare even when agents have no private information. The reason my result is different from Morris & Shin (2002)'s is that in Morris & Shin (2002) decentralization bears no costs: under perfect information, agents agree to coordinate on the same action and welfare is maximized. Conversely, in my model agents' incentives may conflict with social goals: if one agent is a good teammate and the other is a bad teammate, then the good agent does not agree to cooperate; however, if the size of the positive externality is greater than the size of the negative one then it would be socially optimal that they agree to cooperate. When this is the case, welfare is not maximized under perfect information. In fact, in my model, if the size of the positive externality is smaller than the size of the negative externality then drawing a public signal cannot decrease expected social welfare. This follows because in this case, there is no ex post tension between the agents' incentives and the planner's objectives, as in Morris & Shin (2002).

From a theoretical point of view, my model can be seen as a two-person hedonic game (Alcalde & Romero-Medina (2006); Banerjee et al. (2001); Bogomolnaia & Jackson (2002); Dreze & Greenberg (1980)) where there is public uncertainty about the utility of belonging to the grand coalition. Since it features uncertainty about the private returns to cooperation, my model relates the literature on coalitional games with incomplete information (Forges & Serrano (2013); Myerson (1983, 1984, 2005, 2007)). However, my model differs substantially from this literature in that utility is not transferable and, more importantly, it does not feature private information (in my model information is imperfect but symmetric).[4]

My model also relates to the vast literature on teams in information and organizational economics. Team production is common in many organizations (e.g., an interdisciplinary

[4]In fact, given the principal's choice about whether to draw a signal, my model can be written as a coalitional game with public incomplete information, meaning that the agents share the same beliefs about the type profile and the other agents' beliefs even in the interim stage of the game. Following Harsanyi (1967), a player's type is a characterization of his beliefs about the state of the world and the beliefs of the other players. Notice that in my model an agent's state (whether he is a good or bad teammate) does not coincide with his type, in the sense of Harsanyi. Actually, when writing my model as a coalitional game with incomplete information, each agent's type set coincides with the set of all possible realizations of the signal of the agents' states.

board of physicians suggesting surgery to a patient or a group of lawyers offering legal consultancy services). Research in organizational and personnel economics has documented that self-managed teams are becoming increasingly popular in workplaces (Adler (1997); Bandiera et al. (2013); Hamilton et al. (2003); McHugh (1997)). Team production generates externalities among workers because a worker's output in a team depends on whether his peers can help him to fulfill his task (see e.g. Moreland et al. (2002); Mumford et al. (2008); Palanski et al. (2011); Slavin (2011)). I contribute to this literature by analyzing how workers' incentives to team up change when they cannot perfectly observe whether their peers can help them when deciding whether to team up.

Economists have been interested in the design of optimal incentive schemes for teams when team production suffers moral hazard or adverse selection problems (Che & Yoo (2001); Holmstrom (1982); McAfee & McMillan (1991); Rayo (2007)). First, my model differs from this literature because there is no moral hazard or adverse selection issue. Specifically, while information about the private returns to cooperation can be imperfect, it is always symmetric. Second, my model departs from this literature because each worker is assumed to be paid for the output of the task he is assigned to, so that the wage scheme is fixed. Most of the literature on team production under uncertainty (whether arising because effort is unobservable or workers hold private information about payoff-relevant characteristics) assumes that a team is already formed, with a few notable exceptions (Rahman (2005); Tumennasan (2011, 2014); von Siemens & Kosfeld (2014)). I contribute to this literature by studying (i) how public uncertainty about whether a worker is a good or bad teammate affects the worker's incentives to team up, and (ii) the welfare consequences of this uncertainty.

3.2 The Model

Players. Consider a consulting firm composed of a risk neutral manager (principal) and two risk neutral consultants (agents). Let $N := \{1, 2\}$ be an index set for the set of agents.

Project. A client commissions a market analysis report (project) to the firm. The project is composed of two tasks: statistical analysis of market data (task s) and writing of the final draft (task w). Agents 1 and 2 are assigned tasks s and w, respectively. Project difficulty is a random variable $\delta := (\delta_s, \delta_w)$, where δ_s and δ_w are the difficulties of task s and w, respectively. Project difficulty takes values on a commonly known set

$\Delta := \{\delta^1, \delta^2, \delta^3, \delta^4\} \subset \mathbb{R}^2$, and is drawn from a commonly known distribution such that

$$\begin{aligned}\Pr\left(\delta = \delta^1\right) &= p^2\\ \Pr\left(\delta = \delta^2\right) &= p\,(1-p)\\ \Pr\left(\delta = \delta^3\right) &= (1-p)\,p\\ \Pr\left(\delta = \delta^4\right) &= (1-p)^2\,,\end{aligned}$$

for some $p \in (0,1)$. Such distribution is referred to as the common prior belief.

Agents' abilities. For each $i \in N$, agent i's ability is a commonly known vector $\alpha^i := (\alpha^i_s, \alpha^i_w) \in \mathbb{R}^2$, where α^i_s and α^i_w are agent i's statistical knowledge and writing skills, respectively. For each $\tau \in \{s, w\}$, if $\alpha^i_\tau \geq \delta_\tau$ then agent i is said to be good fit to the task τ, and $\alpha^i_\tau < \delta_\tau$ then agent i is said to be bad fit to task τ.

States. If $\alpha^1_w \geq \delta_w$ ($\alpha^1_w < \delta_w$) then agent 1 is said to be good (bad) fit to task w, and if $\alpha^2_s \geq \delta_s$ ($\alpha^2_s < \delta_s$) then agent 2 is said to be good (bad) fit to task s. Assume that[5]

- $\alpha^1_w \geq \delta^1_w$ and $\alpha^2_s \geq \delta^1_s$
- $\alpha^1_w \geq \delta^2_w$ and $\alpha^2_s < \delta^2_s$,
- $\alpha^1_w < \delta^3_w$ and $\alpha^2_s \geq \delta^3_s$,
- $\alpha^1_w < \delta^4_w$ and $\alpha^2_s < \delta^4_s$.

Thus, if project 1 is commissioned then agent 1 is good fit to task w and agent 2 is good fit to task s, if project 2 is commissioned then agent 1 is good fit to task 2 and agent 2 is bad fit to task s, if project 3 is commissioned then agent 1 is bad fit to task w and agent 2 is good fit to task s, and if project 4 is commissioned then agent 1 is bad fit to task w and agent 2 is bad fit to task s.

Let $\Theta^i := \{\beta, \gamma\}$, where β and γ stand for bad and good, respectively, and define a function $\theta := (\theta^1, \theta^2) : \Delta \to \Theta := \Theta^1 \times \Theta^2$ such that

$$\theta(\delta) = \begin{cases} (\beta, \beta) & \text{if } \alpha^1_w < \delta_w \text{ and } \alpha^2_s < \delta_s \\ (\beta, \gamma) & \text{if } \alpha^1_w < \delta_w \text{ and } \alpha^2_s \geq \delta_s \\ (\gamma, \beta) & \text{if } \alpha^1_w \geq \delta_w \text{ and } \alpha^2_s < \delta_s \\ (\gamma, \gamma) & \text{if } \alpha^1_w \geq \delta_w \text{ and } \alpha^2_s \geq \delta_s \end{cases}.$$

[5]This assumption is without loss of generality as long as no state (see below) is redundant.

Then,

$$\begin{aligned}
\Pr(\theta = (\gamma, \gamma)) &= p^2 \\
\Pr(\theta = (\gamma, \beta)) &= p(1-p) \\
\Pr(\theta = (\beta, \gamma)) &= (1-p)\,p \\
\Pr(\theta = (\beta, \beta)) &= (1-p)^2 .
\end{aligned}$$

Refer to Θ^i as agent i's state set and to θ^i as agent i's state. Thus, agent 1's state is γ (β) if he is good (bad) fit to task w and agent 2's state is γ (β) if he is good (bad) fit to task s. The prior belief over Δ induces a prior belief over Θ such that the agents' states are identically and independently distributed, with

$$\begin{aligned}
\Pr\left(\theta^i = \gamma\right) &= \sum_{\theta^j \in \Theta^j} \Pr\left(\left(\theta^i, \theta^j\right) = \left(\gamma, \theta^j\right)\right) \\
&= p(1-p) + p^2 \\
&= p.
\end{aligned}$$

Output function. The agents can work on their own tasks separately or they can agree to (commit to) help each other to perform their own tasks (henceforth, team up). If they do not team up then agent i produces $\overline{Y}^i$. If they team up then agent i produces

$$Y^i\left(\theta^j\right) = \begin{cases} \overline{Y}^i + a & \text{if } \theta^j = \gamma \\ \overline{Y}^i - c & \text{if } \theta^j = \beta \end{cases},$$

where a and c are strictly positive scalars such that $a > c$. Thus, agent 1 (that is, the agent who is assigned task s) raises (lowers) his productivity when he teams up and agent 2 is good (bad) fit at task s, and agent 2 (that is, the agent who is assigned task w) raises (lowers) his productivity when he teams up and agent 1 is good (bad) fit at task w. In other words, if agent i receives help from an agent that is good (bad) fit to the task that i is assigned to, then he (i) can produce more (less) than what he would be able to do alone. For this reason, when $\theta^i = \gamma$ ($\theta^i = \beta$) then agent i is referred to as a good (bad) teammate (for j). Notice that when the agents team up, agent i's output can be additively decomposed into an idiosyncratic part ($\overline{Y}^i$), which is the output that he would have produced by working alone on his task, and an externality (a or $-c$), which depends on whether his teammate is bad or good fit to the task i is assigned. Normalize $\overline{Y}^i$ to zero, without loss of generality. Then, an agent's output measures the productive gain or loss that he obtains when he teams up as a function of his teammate's state.

Monitoring and payment. The principal can costlessly monitor each agent's output (that is, each agent's output is verifiable). However, he cannot verify whether the agents cooperate or not, nor he can enforce cooperation. Let w^i be agent i's wage. Each agent is paid for the output he produces, so that $w^i = Y^i$.[6]

Agents' utility functions. Since the agents are risk neutral, if they do not team up then their utility is 0, and if they team up then agent i's utility is

$$u^i\left(\theta^j\right) = \begin{cases} a & \text{if } \theta^j = \gamma \\ -c & \text{if } \theta^j = \beta \end{cases}.$$

An agent's utility gives the utility gain or loss that he obtains when he teams up as a function of his teammate's state.

Agents' ex ante utility functions. Agent i's ex ante utility from teaming up is

$$\begin{aligned} E^p u^i &= \sum_{\theta^j \in \Theta^j} \Pr\left(\theta^j\right) u^i\left(\theta^j\right) \\ &= pa + (1-p)\, c. \end{aligned}$$

An agent's ex ante utility measures the expected gain or loss from teaming up, given his prior belief over Δ.

Public signal. The principal can provide the agents with better information about project difficulty before they choose whether to team up (for example, he might ask the client to be more specific about the features of the project commissioned). Formally, the principal can draw a publicly observable random variable $d := (d_s, d_w)$ taking values on Δ from a commonly known distributions such that

$$\begin{aligned} \Pr\left(d = \delta^1 \mid \delta = \delta^1\right) &= (1-\varepsilon)^2 \\ \Pr\left(d = \delta^2 \mid \delta = \delta^1\right) &= (1-\varepsilon)\,\varepsilon \\ \Pr\left(d = \delta^3 \mid \delta = \delta^1\right) &= \varepsilon\,(1-\varepsilon) \\ \Pr\left(d = \delta^4 \mid \delta = \delta^1\right) &= \varepsilon^2, \end{aligned}$$

$$\begin{aligned} \Pr\left(d = \delta^1 \mid \delta = \delta^2\right) &= (1-\varepsilon)\,\varepsilon \\ \Pr\left(d = \delta^2 \mid \delta = \delta^2\right) &= (1-\varepsilon)^2 \\ \Pr\left(d = \delta^3 \mid \delta = \delta^2\right) &= \varepsilon^2 \\ \Pr\left(d = \delta^4 \mid \delta = \delta^2\right) &= \varepsilon\,(1-\varepsilon)\,, \end{aligned}$$

[6]Contrary to what happens in the literature on team production with asymmetric information, the principal cannot choose the wage scheme.

$$\begin{aligned}
\Pr\left(d=\delta^1 \mid \delta=\delta^3\right) &= \varepsilon\left(1-\varepsilon\right)\\
\Pr\left(d=\delta^2 \mid \delta=\delta^3\right) &= \varepsilon^2\\
\Pr\left(d=\delta^3 \mid \delta=\delta^3\right) &= \left(1-\varepsilon\right)^2\\
\Pr\left(d=\delta^4 \mid \delta=\delta^3\right) &= \left(1-\varepsilon\right)\varepsilon,
\end{aligned}$$

and

$$\begin{aligned}
\Pr\left(d=\delta^1 \mid \delta=\delta^4\right) &= \varepsilon^2\\
\Pr\left(d=\delta^2 \mid \delta=\delta^4\right) &= \varepsilon\left(1-\varepsilon\right)\\
\Pr\left(d=\delta^3 \mid \delta=\delta^4\right) &= \left(1-\varepsilon\right)\varepsilon\\
\Pr\left(d=\delta^4 \mid \delta=\delta^4\right) &= \left(1-\varepsilon\right)^2,
\end{aligned}$$

for some $\varepsilon \in \left[0, \frac{1}{2}\right]$. In the following, ε is referred to as noise.

Appearances. Let $T^i := \{b, g\}$, where b and g stand for bad teammate and good teammate. Refer to T^i as agent i's appearance set, and let t^i denote an element of T^i. Finally, define a function $t := (t^1, t^2) : \Delta \to T := T^1 \times T^2$ such that

$$t\left(d\right) = \begin{cases} (b,b) & \text{if } \alpha_w^1 < d_w \text{ and } \alpha_s^2 < d_s \\ (b,g) & \text{if } \alpha_w^1 < d_w \text{ and } \alpha_s^2 \geq d_s \\ (g,b) & \text{if } \alpha_w^1 \geq d_w \text{ and } \alpha_s^2 < d_s \\ (g,g) & \text{if } \alpha_w^1 \geq d_w \text{ and } \alpha_s^2 \geq d_s \end{cases}.$$

Then,

$$\begin{aligned}
\Pr\left(t=(g,g) \mid \theta=(\gamma,\gamma)\right) &= \left(1-\varepsilon\right)^2\\
\Pr\left(t=(g,b) \mid \theta=(\gamma,\gamma)\right) &= \left(1-\varepsilon\right)\varepsilon\\
\Pr\left(t=(b,g) \mid \theta=(\gamma,\gamma)\right) &= \varepsilon\left(1-\varepsilon\right)\\
\Pr\left(t=(b,b) \mid \theta=(\gamma,\gamma)\right) &= \varepsilon^2,
\end{aligned}$$

$$\begin{aligned}
\Pr\left(t=(g,g) \mid \theta=(\gamma,\beta)\right) &= \left(1-\varepsilon\right)\varepsilon\\
\Pr\left(t=(g,b) \mid \theta=(\gamma,\beta)\right) &= \left(1-\varepsilon\right)^2\\
\Pr\left(t=(b,g) \mid \theta=(\gamma,\beta)\right) &= \varepsilon^2\\
\Pr\left(t=(b,b) \mid \theta=(\gamma,\beta)\right) &= \varepsilon\left(1-\varepsilon\right),
\end{aligned}$$

$$
\begin{aligned}
\Pr\left(t=(g,g)\mid\theta=(\beta,\gamma)\right) &= \varepsilon\left(1-\varepsilon\right)\\
\Pr\left(t=(g,b)\mid\theta=(\beta,\gamma)\right) &= \varepsilon^2\\
\Pr\left(t=(b,g)\mid\theta=(\beta,\gamma)\right) &= \left(1-\varepsilon\right)^2\\
\Pr\left(t=(b,b)\mid\theta=(\beta,\gamma)\right) &= \left(1-\varepsilon\right)\varepsilon,
\end{aligned}
$$

and

$$
\begin{aligned}
\Pr\left(t=(g,g)\mid\theta=(\beta,\beta)\right) &= \varepsilon^2\\
\Pr\left(t=(g,b)\mid\theta=(\beta,\beta)\right) &= \varepsilon\left(1-\varepsilon\right)\\
\Pr\left(t=(b,g)\mid\theta=(\beta,\beta)\right) &= \left(1-\varepsilon\right)\varepsilon\\
\Pr\left(t=(b,b)\mid\theta=(\beta,\beta)\right) &= \left(1-\varepsilon\right)^2.
\end{aligned}
$$

Thus, drawing a noisy public signal d of δ is equivalent to noisy public signal t of θ such that

$$
\begin{aligned}
\Pr\left(t^i=g\mid\theta^i=\gamma\right) &= \sum_{t^j\in T^j}\Pr\left(\left(t^i,t^j\right)=\left(g,\theta^j\right)\mid\left(\theta^i,\theta^j\right)=\left(\gamma,\theta^j\right)\right)\\
&= \left(1-\varepsilon\right)^2+\left(1-\varepsilon\right)\varepsilon\\
&= 1-\varepsilon,
\end{aligned}
$$

$$
\begin{aligned}
\Pr\left(t^i=b\mid\theta^i=\gamma\right) &= \sum_{t^j\in T^j}\Pr\left(\left(t^i,t^j\right)=\left(b,\theta^j\right)\mid\left(\theta^i,\theta^j\right)=\left(\gamma,\theta^j\right)\right)\\
&= \varepsilon^2+\varepsilon\left(1-\varepsilon\right)\\
&= \varepsilon,
\end{aligned}
$$

$$
\begin{aligned}
\Pr\left(t^i=g\mid\theta^i=\beta\right) &= \sum_{t^j\in T^j}\Pr\left(\left(t^i,t^j\right)=\left(g,\theta^j\right)\mid\left(\theta^i,\theta^j\right)=\left(\beta,\theta^j\right)\right)\\
&= \varepsilon\left(1-\varepsilon\right)+\varepsilon^2\\
&= \varepsilon,
\end{aligned}
$$

and

$$
\begin{aligned}
\Pr\left(t^i=b\mid\theta^i=\beta\right) &= \sum_{t^j\in T^j}\Pr\left(\left(t^i,t^j\right)=\left(b,\theta^j\right)\mid\left(\theta^i,\theta^j\right)=\left(\beta,\theta^j\right)\right)\\
&= \left(1-\varepsilon\right)\varepsilon+\varepsilon^2\\
&= 1-\varepsilon.
\end{aligned}
$$

Notice that $\Pr(t^i \mid \theta^i) = \Pr(t^i \mid \theta)$, since θ^i and θ^j are independent. If $t^i = b$ then agent i is said to appear to be a bad teammate (henceforth, bad), and if $t^i = g$ then agent i is said to appear to be a good teammate (henceforth, good). By Bayes' theorem,

$$\Pr\left(\theta^i = \beta \mid t^i = b\right) = \frac{(1-\varepsilon)(1-p)}{(1-\varepsilon)(1-p)+\varepsilon p} := p_h^\beta,$$

$$\Pr\left(\theta^i = \gamma \mid t^i = g\right) = \frac{(1-\varepsilon)p}{(1-\varepsilon)p+\varepsilon(1-p)} := p_h^\gamma,$$

$$\Pr\left(\theta^i = \beta \mid t^i = g\right) = 1 - \frac{(1-\varepsilon)p}{(1-\varepsilon)p+\varepsilon(1-p)} := p_\ell^\beta = 1 - p_h^\gamma,$$

$$\Pr\left(\theta^i = \gamma \mid t^i = b\right) = 1 - \frac{(1-\varepsilon)(1-p)}{(1-\varepsilon)(1-p)+\varepsilon p} := p_\ell^\gamma = 1 - p_h^\beta.$$

Notice that $p_\ell^\gamma < p < p_h^\gamma$ and $p_\ell^\beta < 1-p < p_h^\beta$, and that p_h^β and p_h^γ are strictly decreasing functions of ε, while p_ℓ^β and p_ℓ^γ are strictly increasing functions of ε.

Agents' interim utility functions. Agent i's interim utility from teaming up is

$$E^{p,\varepsilon}u^i\left(t^j\right) = \sum_{\theta^j \in \Theta^j} \Pr\left(\theta^j \mid t^j\right) u^i\left(\theta^j\right).$$

An agent's interim utility measures the expected gain or loss from teaming up, given his belief over Δ updated based on the information disclosed by an appearance profile. Notice that

$$\begin{aligned}
E^{p,0}u^i((g,g)) &= u^i((\gamma,\gamma)) \\
E^{p,0}u^i((g,b)) &= u^i((\gamma,\beta)) \\
E^{p,0}u^i((b,g)) &= u^i((\beta,\gamma)) \\
E^{p,0}u^i((b,b)) &= u^i((\beta,\beta)),
\end{aligned}$$

and

$$E^{p,\frac{1}{2}}u^i\left(t^j\right) = E^p u^i,$$

for each $t^j \in T^j$.

Team formation. If the principal does not draw an appearance profile then the agents team up if and only if $(E^p u^1, E^p u^2) \geq 0$. On the other hand, if he draws an appearance profile then the agents team up if and only if $\left(E^{p,\varepsilon}u^1\left(t^2\right), E^{p,\varepsilon}u^2\left(t^1\right)\right) \geq 0$.

Social welfare. The principal's preferences are such that for each realization of δ (that is, for each realization of θ), he would like to maximize aggregate output. Thus, if $\max\left\{0, Y^1\left(\theta^2\right) + Y^2\left(\theta^1\right)\right\} = 0$ then the principal does not want the agents to team up, and if $\max\{0, Y^1(\theta^2) + Y^2(\theta^1)\} = Y^1(\theta^2) + Y^2(\theta^1)$ then the principal wants the agents to team up. Since the agents are risk neutral and they are paid for the output he produces, aggregate output is equal to the the sum of the agents' utilities. This means that the principal acts a utilitarian social planner. Thus, in the following, aggregate output is sometimes referred to as social welfare.

Expected social welfare. Let

$$\mathbb{I}^{p,\varepsilon}_{team}(t) = \begin{cases} 1 & \text{if} \left(E^{p,\varepsilon}u^1(t^2), E^{p,\varepsilon}u^2(t^1)\right) \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Expected aggregate output (or expected social welfare) when the principal draws an appearance profile is

$$\begin{aligned} E^{p,\varepsilon}SW &= \sum_{\theta\in\Theta} \Pr(\theta) \sum_{t\in T} \Pr(t \mid \theta)\, \mathbb{I}^{p,\varepsilon}_{team}(t) \sum_{i=1}^{2} u^i\left(\theta^j\right) \\ &= \mathbb{I}^{p,\varepsilon}_{team}(t) \sum_{i=1}^{2} E^{p,\varepsilon}u^i\left(t^j\right). \end{aligned}$$

Expected social welfare when the principal does not draw an appearance profile is

$$\begin{aligned} E^{p,\frac{1}{2}}SW &= \mathbb{I}^{p,\frac{1}{2}}_{team}(t) \sum_{\theta\in\Theta} \Pr(\theta) \sum_{t\in T} \Pr(t \mid \theta) \sum_{i=1}^{2} u^i\left(\theta^j\right) \\ &= \mathbb{I}^{p}_{team} \sum_{i=1}^{2} E^p u^i, \end{aligned}$$

where

$$\mathbb{I}^{p}_{team} = \begin{cases} 1 & \text{if} \left(E^p u^1, E^p u^2\right) \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Gain function. The gain of drawing an appearance profile (henceforth, gain) is a function $G^{p,\cdot} : \left[0, \frac{1}{2}\right] \to \mathbb{R}$ such that

$$G^{p,\varepsilon} = E^{p,\varepsilon}SW - E^{p,\frac{1}{2}}SW,$$

The gain of running an appearance profile is the difference in expected social welfare when the principal draws an appearance profile and when he does not – how much ex ante aggregate output is gained or lost when he draws an appearance profile. The principal draws an appearance profile if and only if its gain is positive.

What happens to the agents' incentives to team up when the principal draws an appearance profile? What happens to expected social welfare? Should the principal draw an appearance profile? Any answers to such questions will clearly depend on the parameters of the model – a, c, p, and ε. In the following, I partition the parameter space of the model by singling out two polar cases, which I refer to cases 1 and 2. In case 1, a, c, and p are such that if the principal does not draw an appearance profile (that is, from an ex ante point of view) then the agents agree to help each other in performing their own tasks. Conversely, in case 2, if the principal does not draw an appearance profile then the agents find it ex ante optimal to work separately. I characterize how gain changes as a function of noise in cases 1 and 2 in turn. The effect of drawing an appearance profile on expected social welfare crucially depends on which of the two cases is at hand.

3.2.1 Case 1: $p \geq \frac{c}{a+c}$

Assume that

$$E^p u^i = pa + (1-p)(-c) \geq 0 \Longleftrightarrow p \geq \frac{c}{a+c}.$$

Claim 3.2.1. *If the principal does not draw an appearance profile then the agents team up.*

Thus, expected social welfare when the principal does not draw an appearance profile is

$$\begin{aligned} E^{p,\frac{1}{2}} SW &= \mathbb{I}^p_{team} \sum_{i=1}^{2} E^p u^i \\ &= 2\left(pa + (1-p)(-c)\right) \\ &= p^2 (2a) + 2\left(p(1-p)(a-c)\right) + (1-p)^2(-2c). \end{aligned}$$

Notice that $E^{p,\frac{1}{2}} SW \geq 0 \Longleftrightarrow p \geq \frac{c}{a+c}$. Ex ante, there is no tension between the agents' incentives and the principal's goals: for each $p \in (0,1)$, the agents team up as long as the planner would like them to team up. This is because if the principal does not draw an appearance profile then expected social welfare is just the sum of the agents' expected utilities. Since the agents are ex ante identical, the expected social welfare of a team is positive as long as agent i's ex ante utility from teaming up is positive.

What happens when the principal draws an appearance profile? To answer this question, assume that he draws an appearance profile. As is clear, the outcome will now depend on noise. To begin with, assume that $\varepsilon = 0$. In this case, drawing an appearance profile is equivalent to revealing the state profile. Since an agent is not willing to team up with a bad agent, the following claim holds:

Claim 3.2.2. *Let $\varepsilon = 0$. If the principal draws an appearance profile, then the agents team up if and only if $t = (g, g)$.*

Next, consider the case in which $\varepsilon \in \left(0, \frac{1}{2}\right)$.

Claim 3.2.3. *For each $\varepsilon \in \left[0, \frac{1}{2}\right)$,*

$$E^p u^i \geq 0 \implies E^{p,\varepsilon} u^i (g) > 0.$$

In words, if agent i is willing to team up when the principal does not draw an appearance profile (that is, at the ex ante stage), then he must be more willing to do so when the principal draws an appearance profile (that is, at the interim stage), provided that agent j appears to be good and $\varepsilon \neq \frac{1}{2}$. Thus, as long as $p \geq \frac{c}{a+c}$, an agent is willing to team up with an agent that appears to be good independently of the noise of the appearance profile. On the other hand, if agent j appears to be bad then agent i's willingness to team up crucially depends on ε. This is shown in the following claim:

Claim 3.2.4. *There exists a unique $\varepsilon^* \in \left(0, \frac{1}{2}\right]$ such that*

$$E^{p,\varepsilon} u^i (b) \begin{cases} < 0, & \text{for each } \varepsilon \in (0, \varepsilon^*) \\ = 0 & \text{if } \varepsilon = \varepsilon^* \\ \geq 0 & \text{for each } \varepsilon \in \left(\varepsilon^*, \frac{1}{2}\right) \end{cases}.$$

In words, if the noise of the appearance profile is sufficiently high then, even if an agent appears to be bad, the expected utility of teaming up with him is positive. However, when noise is sufficiently low, if an agent appears to be bad then the expected utility of teaming up with him is negative. Since an agent agrees to cooperate if and only if the expected utility from teaming up is positive, the following claim immediately follows:

Claim 3.2.5. *If $\varepsilon \in [0, \varepsilon^*)$, then the agents team up if and only if $t = (g, g)$. If $\varepsilon \in \left[\varepsilon^*, \frac{1}{2}\right]$, then the agents team up, for each $t \in T$.*

Claim 3.2.5 shows that (i) if noise is sufficiently small, then the agents team up if and only if they both appear to be good; and (ii) if noise is sufficiently large, then the agents team up independently of the appearance profile. Notice that when $\varepsilon \in \left[\varepsilon^*, \frac{1}{2}\right]$, the information disclosed by the appearance profile is so imprecise that it is irrelevant, in the sense that the agents behave as if $\varepsilon = \frac{1}{2}$. In the following, ε^* is referred to as a threshold, an appearance profile such that $\varepsilon \in \left[\varepsilon^*, \frac{1}{2}\right]$ is said to be an irrelevant appearance profile, and an appearance profile such that $\varepsilon \in [0, \varepsilon^*)$ is said to be an relevant appearance profile.

By setting $E^{p,\varepsilon}u^i(b) = 0$ and solving for ε, one gets

$$\varepsilon^* = \frac{(1-p)\,c}{pa + (1-p)\,c}.$$

Notice that $\varepsilon^* \in [0,1]$, by construction. Moreover, $\varepsilon^* \leq \frac{1}{2}$ as long as $p \geq \frac{c}{a+c}$ (that is, $E^p u^i \geq 0$).

Characterization

In this paragraph, I characterize how gain changes as a function of noise in the case in which $p \geq \frac{c}{a+c}$. The results are summarized in the following statement:

Statement 3.2.1. *Assume $p \geq \frac{c}{a+c}$. The gain of drawing a relevant appearance profile can be negative, and it can be a non-monotonic function of noise. As long as $p \neq \frac{c}{a+c}$, the relationship between gain and noise exhibits a discontinuity at the point in which the appearance profile becomes irrelevant. Finally, as long as $p \geq \frac{c-a}{c-3a}$, gain is strictly decreasing in noise.*

First, consider the case in which the appearance profile is irrelevant (that is, $\varepsilon \in \left[\varepsilon^*, \frac{1}{2}\right]$).

Claim 3.2.6. *For each $\varepsilon \in \left[\varepsilon^*, \frac{1}{2}\right]$, $G^{p,\varepsilon} = 0$.*

Intuitively, if ε reaches threshold ε^* then the agents behave as if the principal did not draw an appearance profile, so that the expected social welfare when the principal draws an appearance profile is equal to expected social welfare when he does not draw it.

The more interesting case is when the appearance profile is relevant (that is, when $\varepsilon \in [0, \varepsilon^*)$). In this case, by Claim 3.2.1, if the principal does not draw an appearance profile then the agents team up whatever the state profile. On the other hand, by Claim 3.2.5, if he draws an appearance profile then the agents team up if and only if $t = (g, g)$. The following lemma outlines a way to write the gain of drawing a relevant appearance profile that turns out to be useful in proving Theorem 3.2.1.

Lemma 3.2.1. *For each $\varepsilon \in [0, \varepsilon^*)$,*

$$\begin{aligned} G^{p,\varepsilon} &= p^2(1-\varepsilon)^2 2a + 2p(1-p)\varepsilon(1-\varepsilon)(a-c) + (1-p)^2\varepsilon^2(-2c) - \\ &\quad p^2(2a) + 2p(1-p)(a-c) + (1-p)^2(-2c). \end{aligned}$$

The following results follow immediately:

Corollary 3.2.1. *For each* $\varepsilon \in [0, \varepsilon^*)$*, if* $p > \frac{1}{2}$ *then* $\frac{\partial^2 G^{p,\varepsilon}}{\partial \varepsilon^2} > 0$*, if* $p = \frac{1}{2}$ *then* $\frac{\partial^2 G^{p,\varepsilon}}{\partial \varepsilon^2} = 0$*, and if* $p < \frac{1}{2}$ *then* $\frac{\partial^2 G^{p,\varepsilon}}{\partial \varepsilon^2} < 0$.

In order to get an intuition for this, notice that, when $\varepsilon \in [0, \varepsilon^*)$, $G^{p,\varepsilon} = \sum_{i=1}^{2} E^{p,\varepsilon} u^i(g) - E^{p,\frac{1}{2}} SW$. Then, recall that

$$\begin{aligned} E^{p,\varepsilon} u^i(g) &= p_h^\gamma a + p_\ell^\beta (-c) \\ &= \left(\frac{(1-\varepsilon) p}{(1-\varepsilon) p + \varepsilon (1-p)} \right) a + \left(1 - \frac{(1-\varepsilon) p}{(1-\varepsilon) p + \varepsilon (1-p)} \right) (-c). \end{aligned}$$

On the one hand, if $p > \frac{1}{2}$, then p_h^γ is strictly convex in ε and p_ℓ^γ is strictly concave in ε; if $p = \frac{1}{2}$, then p_h^γ and p_h^γ are linear ε; and if $p < \frac{1}{2}$, then p_h^γ is strictly concave in ε and p_ℓ^γ is strictly convex in ε. On the other hand, $E^{p,\frac{1}{2}} SW$ does not depend on ε. Thus, the result follows.

Let $f : \mathbb{R} \to \mathbb{R}$ be a function such that

$$\begin{aligned} f(\varepsilon) &= p^2 (1-\varepsilon)^2 2a + 2p(1-p)\varepsilon(1-\varepsilon)(a-c) + (1-p)^2 \varepsilon^2 (-2c) - \\ &\quad p^2 (2a) + 2p(1-p)(a-c) + (1-p)^2 (-2c). \end{aligned}$$

Corollary 3.2.2. *Let* $p \neq \frac{1}{2}$*. Denote the solution to* $\frac{\partial f(\varepsilon)}{\partial \varepsilon} = 0$ *by* $\widetilde{\varepsilon}$*. If* $p < \frac{c-a}{c-3a}$ *then* $\widetilde{\varepsilon} > 0$*; if* $p \in \left[\frac{c-a}{c-3a}, \frac{1}{2}\right)$*, then* $\widetilde{\varepsilon} \leq 0$*; and if* $p > \frac{1}{2}$ *then* $\widetilde{\varepsilon} > \varepsilon^*$.

The rationale for Corollary is 3.2.2 is that, from Lemma 3.2.1, one can see that $G^{p,\varepsilon}$ is a degree 2 polynomial on $[0, \varepsilon^*)$. One may ask where the stationary point of such polynomial falls. This question is interesting because would the stationary point fall in $(0, \varepsilon^*)$, there would be a non-monotonocity in the effect of noise on gain. Corollary 3.2.2 shows that (i) if $p \geq \frac{c-a}{c-3a}$ (that is, if $p \in \left[\frac{c-a}{c-3a}, \frac{1}{2}\right)$ or $p > \frac{1}{2}$) then the stationary point cannot fall in $[0, \varepsilon^*)$, so that $G^{p,\varepsilon}$ must be a monotone function of ε on $[0, \varepsilon^*)$; (ii) while if $p < \frac{c-a}{c-3a}$, then the stationary point may fall in $[0, \varepsilon^*)$, so that $G^{p,\varepsilon}$ could be a non-monotone function of ε on $[0, \varepsilon^*)$. Given that $\frac{c-a}{c-3a} < \frac{1}{2}$, by Corollary 3.2.1, if $\widetilde{\varepsilon} \in (0, \varepsilon^*)$ then it must be the case that, starting at $\varepsilon = 0$, increasing the noise of a relevant signal would initially increase gain, until a maximum is reached at $\widetilde{\varepsilon}$.

Corollary 3.2.3. *If* $p < \frac{c}{a}$ *then* $G^{p,0} > 0$*, if* $p = \frac{c}{a}$ *then* $G^{p,0} = 0$*, and if* $p > \frac{c}{a}$ *then* $G^{p,0} < 0$.

Corollary 3.2.3 shows that when p is sufficiently high (low) then drawing a perfectly informative appearance profile (that is, an appearance profile such that $\varepsilon = 0$) would decrease (increase) expected social welfare relative to the ex ante stage. This is intuitive: when p is sufficiently high, the probability that state (β, β) realizes is so low that the increase in expected aggregate output enjoyed by avoiding the possibility that the agents team up when they are both bad does not compensate the decrease in expected aggregate output suffered because the agents do not team up when one of them is good and the other is bad. The situation is reversed when p is sufficiently low, in which case the probability that (β, β) is so high that the planner prefers avoiding that the agents team up when they are both bad to allowing the possibility that the agents team up when one of them is good and the other is bad.

Corollary 3.2.4. *When* $p = \frac{c}{a+c}$ *then* $\lim_{\eta \to 0} G^{p,\varepsilon^*-\eta} = 0$*, and when* $p > \frac{c}{a+c}$ *then* $\lim_{\eta \to 0} G^{p,\varepsilon^*-\eta} < 0$*.*

Corollary 3.2.4 shows that when $p > \frac{c}{a+c}$, then drawing a sufficiently noisy relevant signal always decreases expected social welfare relative to the ex ante stage.

Theorem 3.2.1 (Characterization of the behavior of $G^{p,\varepsilon}$ as a function of ε)**.** *For each* $\varepsilon \in \left[\varepsilon^*, \frac{1}{2}\right]$*,* $G^{p,\varepsilon} = 0$*. Assume* $p \geq \frac{c-a}{c-3a}$*. For each* $\varepsilon \in [0, \varepsilon^*)$*,* $G^{p,\varepsilon}$ *is smooth and strictly decreasing. Moreover, (i)*

- *if* $p > \frac{c}{a}$ *then* $G^{p,0} < 0$*,*
- *if* $p = \frac{c}{a}$ *then* $G^{p,\varepsilon} = 0$*,*
- *if* $p \in \left(\frac{c}{a+c}, \frac{c}{a}\right)$ *then there exists* $\varepsilon' \in (0, \varepsilon^*)$ *such that* $G^{p,\varepsilon'} = 0$*,*
- *if* $p = \frac{c}{a+c}$ *then* $\lim_{\eta \to 0} G^{p,\varepsilon^*-\eta} = 0$*;*

and (ii)

- *if* $p > \frac{1}{2}$ *then* $G^{p,\varepsilon}$ *is strictly convex on* $[0, \varepsilon^*)$*,*
- *if* $p = \frac{1}{2}$ *then* $G^{p,\varepsilon}$ *is linear on* $[0, \varepsilon^*)$*,*
- *if* $p < \frac{1}{2}$ *then* $G^{p,\varepsilon}$ *is strictly concave on* $[0, \varepsilon^*)$*.*

Assume $p < \frac{c-a}{c-3a}$*. For each* $\varepsilon \in [0, \varepsilon^*)$*,* $G^{p,\varepsilon}$ *is smooth and strictly concave. Moreover, if* $\widetilde{\varepsilon} < \varepsilon^*$ *then* $G^{p,\varepsilon}$ *reaches a maximum at* $\widetilde{\varepsilon}$*.*

Theorem 3.2.1 characterizes how gain changes as a function of noise when p is sufficiently high (that is, when $p \geq \frac{c-a}{c-3a}$). It also shows that when p is sufficiently low (that is, when $p < \frac{c-a}{c-3a}$), then the gain of drawing an appearance profile can be a non-monotonic function of noise. Unfortunately, there is no easy analytical way to characterize how gain changes a function of noise when $p < \frac{c-a}{c-3a}$. However, the qualitative behavior of the gain function can be easily characterized numerically. First of all, one would need to check whether $\widetilde{\varepsilon} < \varepsilon^*$ or $\widetilde{\varepsilon} \geq \varepsilon^*$. In the latter case, the gain of drawing a relevant appearance profile is strictly decreasing in noise, and one just need to check the signs of $G^{p,0}$ and $\lim_{\eta \to 0} G^{p,\varepsilon^*-\eta}$. In the former case, the gain of drawing a relevant appearance profile is a parabola opening downward, and one just need to check the signs of $G^{p,0}$, $G^{p,\widetilde{\varepsilon}}$, and $\lim_{\eta \to 0} G^{p,\varepsilon^*-\eta}$.

3.2.2 Case 2: $p < \frac{c}{a+c}$

Assume that

$$E^p u^i = pa + (1-p)(-c) < 0 \iff p < \frac{c}{a+c}.$$

Claim 3.2.7. *If the principal does not drawn an appearance profile then the agents do not team up.*

Thus, expected social welfare is

$$\begin{aligned} E^{p,\frac{1}{2}} SW &= \mathbb{I}^p_{team} \sum_{i=1}^{2} E^p u_i \\ &= 0. \end{aligned}$$

Notice that, as in Subsection 3.2.1, from an ex ante point of view, there is no tension between the agents' incentives and the principal's objectives: $E^{p,\frac{1}{2}} < 0 \iff p < \frac{c}{a+c}$. Again, the reason is that if the principal does not draw an appearance profile then expected social welfare is equal to the sum of the agents' expected utilities, so that the expected social welfare of a team is negative as long as agent i's ex ante utility from teaming up is negative.

As before, in order to analyze what happens when the principal draws an appearance profile, suppose that he does. Begin with assuming that $\varepsilon = 0$. By Claim 3.2.2, the agents team up if and only if $t = (g, g)$. Next, assume that $\varepsilon \in \left(0, \frac{1}{2}\right)$.

Claim 3.2.8. *For each* $\varepsilon \in \left[0, \frac{1}{2}\right]$,

$$E^p u^i < 0 \implies E^{p,\varepsilon} u^i (b) < 0.$$

Intuitively, if agent i is not willing to team up when the principal does not draw an appearance profile, then he must also be unwilling to do so when the principal draws an appearance profile and agent j appears to be bad. That is, when $p < \frac{c}{a+c}$, an agent is unwilling to team up with an agent that appears to be bad, independently of ε. However, if agent j appears to be good then agent i's willingness to team up is a function of ε, as shown in the following claim.

Claim 3.2.9. *There exists a unique* $\varepsilon^* \in \left(0, \frac{1}{2}\right)$ *such that*

$$E^{p,\varepsilon}u^i(g) \begin{cases} > 0 & \text{for each } \varepsilon \in (0, \varepsilon^*) \\ = 0 & \text{if } \varepsilon = \varepsilon^* \\ < 0 & \text{for each } \varepsilon \in \left(0, \frac{1}{2}\right) \end{cases}.$$

In other words, if the noise of the appearance profile is sufficiently high then, even if an agent appears to be good, the expected utility of teaming up with him is negative. On the other hand, when noise is sufficiently low, if an agent appears to be good then the expected utility of teaming up with him is positive. Since an agent agrees to cooperate if and only if the expected utility from teaming up is positive, the following claim follows:

Claim 3.2.10. *If* $\varepsilon \in [0, \varepsilon^*]$, *then the agents team up if and only if* $t = (g, g)$. *If* $\varepsilon \in \left(\varepsilon^*, \frac{1}{2}\right]$ *then the agents do not team up, for each* $t \in T$.

Claim 3.2.10 shows that (i) if noise is sufficiently small, then the agents team up as long as they both appear to be good; and (ii) if noise is sufficiently large, then the agents do not team up independently of the appearance profile. By setting $E^{p,\varepsilon}u^i(g) = 0$ and solving for ε, one gets

$$\varepsilon^* = \frac{pa}{p(a-c)+c}.$$

Notice that $\varepsilon^* \in [0, 1]$, by construction. Moreover, $\varepsilon^* \leq \frac{1}{2}$ as long as $p < \frac{c}{a+c}$ (that is, $E^p u^i < 0$).

Characterization

In this paragraph, I characterize how gain changes as a function of noise in the case in which $p < \frac{c}{a+c}$. The results are summarized in the following statement:

Statement 3.2.2. *Assume* $p < \frac{c}{a+c}$. *The gain of drawing a relevant appearance profile is always strictly positive, and it continuously strictly decreasing in noise. Moreover, as long as* $p \geq \frac{c}{a+c}$, *gain is strictly decreasing in noise.*

First, consider the case in which the appearance profile is irrelevant.

Claim 3.2.11. *For each* $\varepsilon \in \left(\varepsilon^*, \frac{1}{2}\right]$, $G^{p,\varepsilon} = 0$.

Intuitively, if ε exceeds threshold ε^* then the agents behave as if no appearance profile was drawn, so that expected social welfare when drawing an appearance profile is equal to expected social welfare when not drawing it.

Next, assume that $\varepsilon \in [0, \varepsilon^*]$. In this case, by Claim 3.2.7, if the principal does not draw an appearance profile then the agents do not team up. On the other hand, by Claim 3.2.10, if the he draws an appearance profile then the agents team up if and only if $t = (g, g)$. The following lemma presents a way to write the gain of drawing a relevant appearance profile that turns out to be useful in proving Theorem 3.2.2.

Lemma 3.2.2. *For each* $\varepsilon \in [0, \varepsilon^*)$,

$$G^{p,\varepsilon} = p^2 (1-\varepsilon)^2 2a + 2p(1-p)\varepsilon(1-\varepsilon)(a-c) + (1-p)^2 \varepsilon^2 (-2c).$$

The following lemma is somewhat redundant, in the sense it follows from Theorem 3.2.2. At any rate, for the sake of clarity, I present it here as a separated result.

Lemma 3.2.3. *For each* $\varepsilon \in [0, \varepsilon^*)$, $G^{p,\varepsilon} \geq 0$.

Lemma 3.2.3 shows that the gain of drawing a relevant appearance profile is positive. The intuition is the following: if the principal does not draw an appearance profile then the agents do not team up, so that expected social welfare is 0. On the other hand, if the principal draws a relevant appearance profile then the agents team up as long as $t = (g, g)$, so that expected social welfare is equal to the sum of the agents' interim utilities from teaming up when $t = (g, g)$. Since the agents team up as long as their interim utility from doing so is positive, so must be the gain of drawing a relevant appearance profile. The following results follow:

Corollary 3.2.5. *For each* $\varepsilon \in [0, \varepsilon^*)$, $\frac{\partial^2 G^{p,\varepsilon}}{\partial \varepsilon^2} < 0$.

An intuition for this result follows immediately from the intuition for Corollary 3.2.1.

Let $f : \mathbb{R} \to \mathbb{R}$ be a function such that

$$f(\varepsilon) = p^2 (1-\varepsilon)^2 2a + 2p(1-p)\varepsilon(1-\varepsilon)(a-c) + (1-p)^2 \varepsilon^2 (-2c).$$

Corollary 3.2.6. *Denote the solution to* $\frac{\partial f(\varepsilon)}{\partial \varepsilon} = 0$ *by* $\widetilde{\varepsilon}$. *If* $p < \frac{c-a}{c-3a}$ *then* $\widetilde{\varepsilon} > 0$; *if* $p \in \left[\frac{c-a}{c-3a}, \frac{1}{2}\right)$, *then* $\widetilde{\varepsilon} \leq 0$.

The rationale for Corollary 3.2.6 is the same as the rationale for Corollary 3.2.2.

Corollary 3.2.7. $G^{p,0} > 0$.

Corollary 3.2.7 shows that the gain of drawing a perfectly informative appearance profile is always strictly positive. This is clear: if the principal draws a perfectly informative appearance profile then the agents team up as long as they are both good (which happens with probability p^2), and in this case aggregate output is equal to $2a$. On the other hand, if he does not draw an appearance profile then aggregate output is equal to 0. In other words, by drawing a perfectly informative appearance profile, the principal can incentivize the agents to confer externalities on each other at least in the case in which $\theta = (\gamma, \gamma)$, while if he does not draw an appearance profile then no externalities are unleashed.

Corollary 3.2.8. $\lim_{\eta \to 0} G^{p,\varepsilon^* - \eta} = 0$.

Corollary 3.2.8 shows that drawing a sufficiently noisy relevant appearance profile always drive gain to 0. This is clear, because $\lim_{\eta \to 0} E^{p,\varepsilon^* - \eta} u^i(g) = 0$, and, by Lemma 3.2.3, $G^{p,\varepsilon} = \Pr(t = (g,g)) \left(\sum_{i=1}^{2} E^{p,\varepsilon} u^i(g) \right)$ on $[0, \varepsilon^*)$.

Theorem 3.2.2 (Characterization of the behavior of $G^{p,\varepsilon}$ as a function of ε)**.** *For each $\varepsilon \in \left[\varepsilon^*, \frac{1}{2}\right]$, $G^{p,\varepsilon} = 0$. Assume $p \geq \frac{c-a}{c-3a}$. For each $\varepsilon \in [0, \varepsilon^*)$, $G^{p,\varepsilon}$ is smooth, strictly decreasing, and strictly positive. Moreover, $\lim_{\eta \to 0} G^{p,\varepsilon^* - \eta} = 0$.*
Assume $p < \frac{c-a}{c-3a}$. For each $\varepsilon \in [0, \varepsilon^*)$, $G^{p,\varepsilon}$ is smooth, strictly concave, and strictly positive. Moreover, $\lim_{\eta \to 0} G^{p,\varepsilon^* - \eta} = 0$, and if $\widetilde{\varepsilon} < \varepsilon^*$ then $G^{p,\varepsilon}$ reaches a maximum at $\widetilde{\varepsilon}$.

Theorem 3.2.2 characterizes how gain changes as a function of noise when p is sufficiently high ($p \geq \frac{c-a}{c-3a}$). Moreover, it shows that when p is sufficiently low ($p < \frac{c-a}{c-3a}$), then the gain of drawing an appearance profile can be a non-monotonic function of noise. Again, there is no easy analytical way to characterize how gain changes a function of noise when $p < \frac{c-a}{c-3a}$, but numerically characterizing the qualitative behavior of the gain function is straightforward. First, one needs to check whether $\widetilde{\varepsilon} < \varepsilon^*$ or $\widetilde{\varepsilon} \geq \varepsilon^*$. In the latter case, the gain of drawing a relevant appearance profile is strictly decreasing in noise. In the former case, the gain of drawing a relevant appearance profile is a parabola opening downward.

3.2.3 Discussion

In this subsection, I discuss the results obtained in Section 3.2, as summarized in Statements 3.2.1 and 3.2.2.

In the case outlined in Subsection 3.2.1, providing the agents with more precise information about the private returns to cooperation has three effects. Firstly, the agents team up when they both appear to be good. This is both individually rational and socially desirable. Secondly, if the agents have sufficiently precise information about the private returns to cooperation, then they do not team up when one of them appears to be bad. This is individually rational (if the agent appearing to be good were to accept the help of the agent appearing to be bad, then he would suffer an expected decrease in the output of the task he is assigned to), but it can be socially undesirable, because it does not allow the bad agent to enjoy the expected positive externality that the agent appearing to be good would confer on him. Finally, if the agents have sufficiently precise information about the private returns to cooperation, then they do not agree to cooperate when both of them appear to be bad. This is individually rational and socially desirable. Whether drawing an appearance profile increases or decreases expected social welfare depends on which of these three effects dominate: if the first and third effects dominate the second one then providing the agents with more precise information about the private returns to cooperation increases expected social welfare, while if the second effects dominates the first and third ones then drawing an appearance profile could decrease expected social welfare.

On the other hand, in the case outlined in Subsection 3.2.2, providing the agents with more precise information about the private returns to cooperation unambiguously increases expected social welfare. The reason why this is the case is that if the agents are left with their prior beliefs then they do not team up, while drawing a relevant appearance profile incentivize them to team up at least in the case in which they both appear to be good.

Notice that in both cases 1 and 2, there exists a point $\varepsilon^* \in \left(0, \frac{1}{2}\right)$ such that if $\varepsilon > \varepsilon^*$, then providing the agents with more precise information about the utility of teaming up does not affect their incentives (and, consequently, expected social welfare). This is precisely the point at which there is a discontinuity in the gain function.

In the case outlined in Subsection 3.2.1, if the principal draws an appearance profile then the agents team up independently of their states. This can be preferable to a situation in which the agents team up as long as they appear to be good. The reason why this happens is the following: even though drawing a relevant appearance profile decreases the chances that the agents team up when they are both bad (which is an undesirable situation, both from an individual and a social point of view), drawing a relevant signal decreases the chances that the agents team up when one of them is good and the other is bad. If the expected costs of increasing the probability that the agents do not team up when $t \in \{(\beta, \gamma), (\gamma, \beta)\}$ outweigh the expected benefits of decreasing the probability that they team up when $t = (\beta, \beta)$, then drawing an appearance profile decreases expected social welfare. Subsection 3.2.1 also shows that there cases in which the principal finds it optimal to provide the agents with more precise information about the private returns to cooperation. This happens when the prior probability that an agent is good (p) is small. An intuition for this is as follows: when the principal does not draw an appearance profile, if p decreases, then the probability that the agents team up when they are all bad increases. Moreover, as long as $p < \frac{1}{2}$, a decrease in p also decreases the probability that the agents team up when one of them is bad and the other is good.[7] All of this decreases the ex ante appealingness of not drawing an appearance profile.

In the case outlined in Subsection 3.2.2, providing the agents with better information about the private returns to cooperation is intrinsically good, in the sense that, from an ex ante point of view, the principal always prefers to draw a relevant appearance profile. Here, the intuition is that if the agents do not have any information (besides their prior beliefs), then no team would be formed, independently of their states. Drawing a relevant appearance profile must be optimal, because then the agents would team up at least when it is very likely that they are both good.

Finally, in both cases 1 and 2, if p is sufficiently low (that is, when $p < \frac{c-a}{c-3a}$), then the gain of drawing a relevant appearance profile can be a non-monotonic function of noise. In this sense, the effect of increasing the precision of the appearance profile on the expected social welfare obtained when drawing can be non-trivial.

3.3 Conclusions

Agents team up to confer externalities on each other. If there is public uncertainty about whether an agent is a good or bad teammate, then there is public uncertainty about the private returns to cooperation. In this chapter I outline a simple model of team formation under uncertainty about the private returns to cooperation, and I show that there are cases in which *more precise information about the private returns to cooperation is worse*, in the sense that, from an ex ante point of view, a utilitarian social planner would choose *not* to provide such more precise information. In this sense, there are cases in which it is "better not to know."

Providing the agents with more precise information about the private returns of cooperation can increase the tension between the agents' incentives and the principal's objectives, because the expected decrease in aggregate output derived from the fact that an agent appearing to be a good teammate does not agree to cooperate with an agent appearing to be a bad teammate can outweigh the expected increase in aggregate output derived from the fact that the agents team up when they both appear to be good and do not team up when they both appear to be bad.

[7] To see this, notice that the probability that $t \in \{(\beta, \gamma), (\gamma, \beta)\}$ is $2p(1-p)$. Clearly, $2p(1-p)$ is decreasing in p as long as $p \leq \frac{1}{2}$.

I am led by two motivations. The first one is to understand the welfare effects of public uncertainty about the private returns of cooperation in an environment in which agents can agree to form groups to confer externalities on each other. One often listen the argument that higher transparency in the banking sector (in the sense of a public disclosure of information about banks' risk exposure, for example) would be socially desirable (see, for instance, Cordella & Yeyati (1998); Nier (2005); Tadesse (2006)). In fact, the transparency of the banking sector is an issue of utmost importance in recommendations on banking laws and regulations such as Basel II and IAS-IFRS. Even though higher transparency can certainly have beneficial effects, in that it can alleviate banks' moral hazard problems (Y. Chen & Hasan (2006); Klein et al. (2013)), economists have largely ignored the fact that information has another fundamental effect on banks' behavior: it affects the way in which banks connect through financial interdependencies. Higher transparency incentivizes "good" banks to group together and segregate out from "bad" banks, which would be left alone. If segregation is disruptive for the liquidity of the banking sector, more transparency can be harmful.

The second one is that there is a growing literature in economics studying the organization of agents into groups and networks. This literature claims that networks plays a fundamental role in shaping the outcomes of many social interactions. Part of this literature is concerned with the formation of social structure (see B. Dutta & Jackson (2013); Hajduková (2006); Jackson (2005), for example). Many models of group and network formation can be conceived as games in which agents deliberately get together to confer externalities on each other. In most of these models information is assumed to be perfect. My model is a first attempt to study the incentive and welfare effects of public uncertainty about the private returns to cooperation in a simple group formation game.

My model could be extended in many directions. In my model, if the planner chooses to draw an appearance profile signal then the information disclosed by the signal received is automatically communicated to the agents. In other words, the principal has to choose whether or not to find out more about the state of the world taking into account that the information revealed by the appearance profile is directly observed by the agents. The story would be different if the principal could draw an appearance profile, update his beliefs based on the information disclosed by the signal received, and then decide whether to reveal this information to the agents. If the agents can observe the principal's choice, this would create a signaling game, in which the agents can infer something about the state of the world from the fact that the principal chooses to draw an appearance profile but does not reveal the information disclosed by the signal realized.

Another possibility is to micro-found the externalities that the agents confer to each other when they team up. For example, one think of a model in which, when the agents team up, they have to simultaneously choose effort levels, and in which a good teammate has a lower cost of providing effort than a bad teammate. Suppose that an agent's effort is a strategic complement to his teammate's effort. If the equilibrium effort provided by a bad teammate is sufficiently low, then the output of the task assigned to an agent when he cooperate with a bad teammate could be lower than the one he would choose when working alone. Similarly, if the equilibrium effort provided by a good teammate is sufficiently high, then the output of the task assigned to an agent when he cooperate with a good teammate could be higher than the one he would choose when working alone. This would provide a strategic micro-foundation to the positive or negative externalities that an agent obtains when teaming up with a good or bad teammate, respectively.

Furthermore, one could assume that the principal can choose the agents' wage schedule, to try to understand what would be an optimal salary scheme.

Bibliography

Ábrahám, Á., & Pavoni, N. (2005). The Efficient Allocation of Consumption under Moral Hazard and Hidden Access to the Credit Market. *Journal of the European Economic Association*, *3*(2-3), 370–381.

Acemoglu, D. (2010). When Does Labor Scarcity Encourage Innovation? *Journal of Political Economy*, *118*(6), 1037–1078.

Adler, P. S. (1997). Work Organization: From Taylorism to Teamwork. *Perspectives on Work*, 61–65.

Advani, A. (2017). *Insurance Networks and Poverty Traps* (Working paper). University College London.

Aker, J. C. (2010). Information from Markets near and Far: Mobile Phones and Agricultural Markets in Niger. *American Economic Journal: Applied Economics*, *2*(3), 46–59.

Alcalde, J., & Romero-Medina, A. (2006). Coalition Formation and Stability. *Social Choice and Welfare*, 365–375.

Algan, Y., & Cahuc, P. (2013). Trust, Institutions and Economic Development. In *Handbook of economic growth* (Vol. 1A). North-Holland.

Ambrus, A., Gao, W. Y., & Milán, P. (2017). *Informal Risk Sharing with Local Information* (Working paper).

Arcand, J.-L., Ai, C., & Éthier, F. (2007). Moral Hazard and Marshallian Inefficiency: Evidence from Tunisia. *Journal of Development Economics*, *83*(2), 411–445.

Arrow, K. J. (1972). Gifts and Exchanges. *Philosophy & Public Affairs*, *1*(4), 343–362.

Attanasio, O. P., & Pavoni, N. (2011). Risk Sharing in Private Information Models with Asset Accumulation: Explaining the Excess Smoothness of Consumption. *Econometrica*, *79*(4), 1027–1068.

Bandiera, O., Barankay, I., & Rasul, I. (2013). Team Incentives: Evidence from a Firm Level Experiment. *Journal of the European Economic Association*, 1079–1114.

Banerjee, S., Konishi, H., & Sönmez, T. (2001). Core in a Simple Coalition Formation Game. *Social Choice and Welfare*, 135–153.

Beaman, L., Karlan, D., Thuysbaert, B., & Udry, C. (2013). Profitability of Fertilizer: Experimental Evidence from Female Rice Farmers in Mali. *American Economic Review*, *103*(3), 381–86.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, Reciprocity, and Social History. *Games and Economic Behavior*, *10*(1), 122–142.

Binswanger, H. P. (1981). Attitudes Toward Risk: Theoretical Implications of an Experiment in Rural India. *The Economic Journal*, *91*(364), 867–890.

Bogomolnaia, A., & Jackson, M. O. (2002). The Stability of Hedonic Coalition Structures. *Games and Economic Behavior*, 201–230.

Bold, T., & Broer, T. (2016). *Risk-Sharing in Village Economies Revisited* (Working paper).

Bold, T., Kaizzi, K. C., Svensson, J., & Yanagizawa-Drott, D. (2017). Lemon Technologies and Adoption: Measurement, Theory and Evidence from Agricultural Markets in Uganda. *The Quarterly Journal of Economics*, *132*(3), 1055–1100.

Bolton, P., & Dewatripont, M. (2005). *Contract Theory.* MIT press.

Bowles, S., & Gintis, H. (2002). Social Capital and Community Governance. *The Economic Journal*, *112*(483), F419–F436.

Branas-Garza, P., Cobo-Reyes, R., Espinosa, M. P., Jiménez, N., Kovářík, J., & Ponti, G. (2010). Altruism and Social Integration. *Games and Economic Behavior*, *69*(2), 249–257.

Budish, E., Che, Y.-K., Kojima, F., & Milgrom, P. (2013). Designing Random Allocation Mechanisms: Theory and Applications. *The American Economic Review*, 585–623.

Carpenter, J. P., Daniere, A. G., & Takahashi, L. M. (2004). Cooperation, Trust, and Social Capital in Southeast Asian Urban Slums. *Journal of Economic Behavior & Organization*, *55*(4), 533–551.

Carpiano, R. M., & Fitterer, L. M. (2014). Questions of Trust in Health Research on Social Capital: What Aspects of Personal Network Social Capital Do They Measure? *Social Science & Medicine*, *116*, 225–234.

Carrell, S. E., Sacerdote, B. I., & West, J. E. (2013). From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation. *Econometrica*, *81*(3), 855–882.

Carroll, G. (2015). Robustness and Linear Contracts. *American Economic Review*, *105*(2), 536–63.

Chandrasekhar, A. (2016). Econometrics of Network Formation. In *The oxford handbook of the economics of networks* (pp. 303–357). Oxford University Press.

Chandrasekhar, A. G., & Jackson, M. O. (2014). *Tractable and Consistent Random Graph Models* (Tech. Rep.). National Bureau of Economic Research.

Che, Y.-K., & Yoo, S.-W. (2001). Optimal Incentives for Teams. *American Economic Review*, 525–541.

Chen, C., Restuccia, D., & Santaeulàlia-Llopis, R. (2017). *The Effects of Land Markets on Resource Allocation and Agricultural Productivity* (Working paper). National Bureau of Economic Research.

Chen, Y., & Hasan, I. (2006). The Transparency of the Banking System and the Efficiency of Information-Based Bank Runs. *Journal of Financial Intermediation*, 307–331.

Cillessen, A. H., & Mayeux, L. (2004). From Censure to Reinforcement: Developmental Changes in the Association Between Aggression and Social Status. *Child Development*, *75*(1), 147–163.

Clark, D. E., Herrin, W. E., Knapp, T. A., & White, N. E. (2003). Migration and Implicit Amenity Markets: Does Incomplete Compensation Matter? *Journal of Economic Geography*, *3*(3), 289–307.

Cochrane, J. H. (1991). A Simple Test of Consumption Insurance. *Journal of Political Economy*, *99*(5), 957–976.

Conlon, J. R. (2009). Two New Conditions Supporting the First-Order Approach to Multisignal Principal–Agent Problems. *Econometrica*, *77*(1), 249–278.

Cordella, T., & Yeyati, E. L. (1998). Public Disclosure and Bank Failures. *International Monetary Fund Staff Papers*, 110–131.

Demange, G., & Wooders, M. (2005). *Group Formation in Economics: Networks, Clubs, and Coalitions*. Cambridge University Press.

Dercon, S., & Christiaensen, L. (2011). Consumption Risk, Technology Adoption and Poverty Traps: Evidence from Ethiopia. *Journal of Development Economics*, *96*(2), 159–173.

Dorfman, J. H. (1996). Modeling Multiple Adoption Decisions in a Joint Framework. *American Journal of Agricultural Economics*, *78*(3), 547–557.

Dreze, J., & Greenberg, J. (1980). Hedonic Coalitions: Optimality and Stability. *Econometrica*, 987–1003.

Duflo, E., Kremer, M., & Robinson, J. (2011). Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya. *The American Economic Review*, *101*(6), 2350–2390.

Dutta, B., & Jackson, M. O. (2013). *Networks and Groups: Models of Strategic Formation*. Springer Science & Business Media.

Dutta, J., & Prasad, K. (2002). Stable Risk-Sharing. *Journal of Mathematical Economics*, *38*(4), 411–439.

Edmans, A., & Gabaix, X. (2011). Tractability in Incentive Contracting. *The Review of Financial Studies*, *24*(9), 2865–2894.

Edward, B. (1958). *The Moral Basis of a Backward Society*. Free Press.

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

Fafchamps, M. (2011). Risk Sharing Between Households. In *Handbook of Social Economics* (pp. 1255–1279). Elsevier.

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

Falk, A., Becker, A., Dohmen, T., Huffman, D., & Sunde, U. (2016). *The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences* (Working paper).

Feder, G., Just, R. E., & Zilberman, D. (1985). Adoption of Agricultural Innovations in Developing Countries: A Survey. *Economic Development and Cultural Change*, 255–298.

Fehr, E. (2009). On the Economics and Biology of Trust. *Journal of the European Economic Association*, *7*(2-3), 235–266.

Fehr, E., Fischbacher, U., & Kosfeld, M. (2005). Neuroeconomic Foundations of Trust and Social Preferences: Initial Evidence. *American Economic Review*, *95*(2), 346–351.

Fehr, E., Fischbacher, U., Von Rosenbladt, B., Schupp, J., & Wagner, G. G. (2003). *A Nation-Wide Laboratory: Examining Trust and Trustworthiness by Integrating Behavioral Experiments into Representative Survey* (Working paper).

Finseraas, H., Hanson, T., Johnsen, Å. A., Kotsadam, A., & Torsvik, G. (2019). *Trust, Ethnic Diversity, and Personal Contact: A Field Experiment* (Working paper).

Forges, F., & Serrano, R. (2013). Cooperative Games with Incomplete Information: Some Open Problems. *International Game Theory Review*, 1340009.

Foster, A. D., & Rosenzweig, M. R. (2001). Imperfect Commitment, Altruism, and the Family: Evidence from Transfer Behavior in Low-Income Rural Areas. *Review of Economics and Statistics*, *83*(3), 389–407.

Foster, A. D., & Rosenzweig, M. R. (2009). *Should Poor Farmers Use More Inputs?* (Working paper).

Foster, A. D., & Rosenzweig, M. R. (2010). Microeconomics of Technology Adoption. *Annual Review of Economics*, *2*(1), 395–424.

Foster, A. D., & Rosenzweig, M. R. (2011). *Are Indian Farms Too Small? Mechanization, Agency Costs, and Farm Efficiency* (Working paper).

Franken, A., Harakeh, Z., Veenstra, R., Vollebergh, W., & Dijkstra, J. K. (2017). Social Status of Adolescents with an Early Onset of Externalizing Behavior: the SNARE Study. *The Journal of Early Adolescence*, *37*(8), 1037–1053.

Fukuyama, F. (1995). *Trust: The Social Virtues and the Creation of Prosperity.* Free Press.

Giné, X., Jakiela, P., Karlan, D., & Morduch, J. (2010). Microfinance Games. *American Economic Journal: Applied Economics*, *2*(3), 60–95.

Giné, X., & Mansuri, G. (2018). Together We Will: Experimental Evidence on Female Voting Behavior in Pakistan. *American Economic Journal: Applied Economics*, *10*(1), 207–35.

Giné, X., & Yang, D. (2009). Insurance, Credit, and Technology Adoption: Field Experimental Evidence from Malawi. *Journal of Development Economics*, *89*(1), 1–11.

Glaeser, E. L., Laibson, D., & Sacerdote, B. (2002). An Economic Approach to Social Capital. *The Economic Journal*, *112*(483), F437–F458.

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

Glanville, J. L., & Paxton, P. (2007). How Do We Learn to Trust? A Confirmatory Tetrad Analysis of the Sources of Generalized Trust. *Social Psychology Quarterly*, *70*(3), 230–242.

Goeree, J. K., McConnell, M. A., Mitchell, T., Tromp, T., & Yariv, L. (2010). The $1/d$ Law of Giving. *American Economic Journal: Microeconomics*, *2*(1), 183–203.

Good, P. (2013). *Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses*. Springer.

Government of India, Ministry of Finance. (2016). *Economic survey* (Technical report).

Guiso, L., Sapienza, P., & Zingales, L. (2009). Cultural Biases in Economic Exchange? *The Quarterly Journal of Economics*, *124*(3), 1095–1131.

Haider, H., Smale, M., & Theriault, V. (2018). Intensification and Intrahousehold Decisions: Fertilizer Adoption in Burkina Faso. *World Development*.

Hajduková, J. (2006). Coalition Formation Games: A Survey. *International Game Theory Review*, 613–641.

Hamilton, B. H., Nickerson, J. A., & Owan, H. (2003). Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation. *Journal of Political Economy*, 465–497.

Harsanyi, J. C. (1967). Games with Incomplete Information Played by "Bayesian" Players. *Management Science*, 1804–1817.

Hawke, S., & Rieger, E. (2013). Popularity, Likeability, and Risk-Taking in Middle Adolescence. *Health*, *5*(06), 41.

Henrich, J. P., Boyd, R., Bowles, S., Fehr, E., Camerer, C., & Gintis, H. (2004). *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies.* Oxford University Press on Demand.

Hirshleifer, J. (1971). The Private and Social Value of Information and the Reward to Inventive Activity. *The American Economic Review*, 561–574.

Holmstrom, B. (1982). Moral Hazard in Teams. *The Bell Journal of Economics*, 324–340.

Holmström, B., & Milgrom, P. (1987). Aggregation and Linearity in the Provision of Intertemporal Incentives. *Econometrica*, 303–328.

Hornbeck, R., & Naidu, S. (2014). When the Levee Breaks: Black Migration and Economic Development in the American South. *The American Economic Review*, *104*(3), 963–990.

Hubert, L., & Schultz, J. (1976). Quadratic Assignment as a General Data Analysis Strategy. *British Journal of Mathematical and Statistical Psychology*, *29*(2), 190–241.

Igarashi, T., Kashima, Y., Kashima, E. S., Farsides, T., Kim, U., Strack, F., . . . Yuki, M. (2008). Culture, Trust, and Social Networks. *Asian Journal of Social Psychology*, *11*(1), 88–101.

Indian Labour Bureau. (2010). *Wage Rates in Rural India* (Tech. Rep.). Labor Bureau, Ministry of Labor & Employment, Government of India.

Inglehart, R. (1997). *Modernization and Post-Modernization: Cultural, Economic and Political Change in 43 Societies.* Princeton University Press.

Jack, B. K. (2013). *Market Inefficiencies and the Adoption of Agricultural Technologies in Developing Countries* (Working paper). Agricultural Technology Adoption Initiative.

Jackson, M. O. (2005). A Survey of Network Formation Models: Stability and Efficiency. In *Group formation in economics: Networks, clubs, and coalitions* (pp. 11–49). Cambridge University Press.

Jackson, M. O. (2018). *A Typology of Social Capital and Associated Network Measures* (Working paper).

Jaga, P., & Patel, Y. (2012). An Overview of Fertilizers Consumption in India: Determinants and Outlook for 2020-A review. *International Journal of Scientific Engineering and Technology*, *1*(6), 285–291.

Jakiela, P., & Ozier, O. (2015). Does Africa Need a Rotten Kin Theorem? Experimental Evidence from Village Economies. *The Review of Economic Studies*, *83*(1), 231–268.

Jalan, J., & Ravallion, M. (1999). Are the Poor Less Well Insured? Evidence on Vulnerability to Income Risk in Rural China. *Journal of Development Economics*, *58*(1), 61–81.

Jewitt, I. (1988). Justifying the First-Order Approach to Principal-Agent Problems. *Econometrica*, *56*(5), 1177–1190.

Just, R. E., & Pope, R. D. (1979). Production Function Estimation and Related Risk Considerations. *American Journal of Agricultural Economics*, *61*(2), 276–284.

Kandel, D. B. (1978). Homophily, Selection, and Socialization in Adolescent Friendships. *American Journal of Sociology*, *84*(2), 427–436.

Kaplan, G. (2006). *The Cross-Sectional Implications of Incomplete Markets: Private Information or Limited Enforcement?* (Working paper). New York University.

Karaivanov, A., & Townsend, R. M. (2014). Dynamic Financial Constraints: Distinguishing Mechanism Design from Exogenously Incomplete Regimes. *Econometrica*, *82*(3), 887–959.

Kawachi, I. (2010). Social Capital and Health. In *Social capital and health* (pp. 1–26). Springer.

Kinnan, C. (2017). *Distinguishing Barriers to Insurance in Thai Villages* (Working paper). Northwestern University.

Klein, T. J., Lambertz, C., & Stahl, K. O. (2013). Market Transparency, Adverse Selection, and Moral Hazard.

Knack, S., & Keefer, P. (1997). Does Social Capital Have an Economic Payoff? A Cross-Country Investigation. *The Quarterly Journal of Economics*, *112*(4), 1251–1288.

Kocherlakota, N., & Pistaferri, L. (2009). Asset Pricing Implications of Pareto Optimality with Private Information. *Journal of Political Economy*, *117*(3), 555–590.

Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., & Fehr, E. (2005). Oxytocin Increases Trust in Humans. *Nature*, *435*(7042), 673.

Kovářík, J., Brañas-Garza, P., Cobo-Reyes, R., Espinosa, M. P., Jiménez, N., & Ponti, G. (2012). Prosocial Norms and Degree Heterogeneity in Social Networks. *Physica A: Statistical Mechanics and its Applications*, *391*(3), 849–853.

Krackardt, D. (1987). QAP Partialling as a Test of Spuriousness. *Social Networks*, *9*(2), 171–186.

Krackhardt, D. (1988). Predicting with Networks: Nonparametric Multiple Regression Analysis of Dyadic Data. *Social Networks*, *10*(4), 359–381.

Laczó, S. (2015). Risk Sharing with Limited Commitment and Preference Heterogeneity: Structural Estimation and Testing. *Journal of the European Economic Association*, *13*(2), 265–292.

Laffont, J.-J., & Matoussi, M. S. (1995). Moral Hazard, Financial Constraints and Sharecropping in El Oulja. *The Review of Economic Studies*, *62*(3), 381–399.

LaFontana, K. M., & Cillessen, A. H. (2002). Children's Perceptions of Popular and Unpopular Peers: A Multimethod Assessment. *Developmental Psychology*, *38*(5), 635.

Leider, S., Möbius, M. M., Rosenblat, T., & Do, Q.-A. (2009). Directed Altruism and Enforced Reciprocity in Social Networks. *The Quarterly Journal of Economics*, *124*(4), 1815–1851.

Ligon, E. (1998). Risk Sharing and Information in Village Economies. *The Review of Economic Studies*, *67*(4), 847–864.

Ligon, E., Thomas, J. P., & Worrall, T. (2002). Informal Insurance Arrangements with Limited Commitment: Theory and Evidence from Village Economies. *The Review of Economic Studies*, *69*(1), 209–244.

Mace, B. J. (1991). Full Insurance in the Presence of Aggregate Uncertainty. *Journal of Political Economy*, *99*(5), 928–956.

Marmaros, D., & Sacerdote, B. (2006). How Do Friendships Form? *The Quarterly Journal of Economics*, *121*(1), 79–119.

Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory.* Oxford University Press.

Mayer, A., & Puller, S. L. (2008). The Old Boy (and Girl) Network: Social Network Formation on University Campuses. *Journal of Public Economics*, *92*(1-2), 329–347.

Mazur, K. (2018). *Sharing Risk to Avoid Tragedy: Theory and Application to Farming* (Working paper). European University Institute.

Mazzocco, M., & Saini, S. (2012). Testing Efficient Risk Sharing with Heterogeneous Risk Preferences. *American Economic Review*, *102*(1), 428–468.

McAfee, P., & McMillan, J. (1991). Optimal Contracts for Teams. *International Economic Review*, 561–577.

McHugh, P. P. (1997). Team-based Work Systems: Lessons from the Industrial Relations Literature. *People and Strategy*, 44.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual review of sociology*, *27*(1), 415–444.

Milán, P. (2016). *Network-Constrained Risk Sharing in Village Economies* (Working paper). Universitat Autónoma de Barcelona and Barcelona GSE.

Miller, D., & Paulson, A. (2007). *Risk Taking and the Quality of Informal Insurance: Gambling and Remittances in Thailand* (Working paper). Federal Reserve Bank of Chicago.

Moreland, R. L., Argote, L., & Krishnan, R. (2002). Training People to Work in Groups. In *Theory and Research on Small Groups* (pp. 37–60). Springer.

Morris, S., & Shin, H. S. (2002). Social Value of Public Information. *The American Economic Review*, 1521–1534.

Morten, M. (2017). Temporary Migration and Endogenous Risk Sharing in Village India. *Journal of Political Economy*. (Forthcoming)

Mouw, T. (2006). Estimating the causal effect of social capital: A review of recent research. *Annu. Rev. Sociol.*, *32*, 79–102.

Mumford, M. D., Hunter, S. T., & Bedell-Avers, K. E. (2008). *Multi-Level Issues in Creativity and Innovation* (Vol. 7). Emerald Group Publishing.

Munshi, K., & Rosenzweig, M. (2006). Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy. *The American Economic Review*, *96*(4), 1225–1252.

Munshi, K., & Rosenzweig, M. (2009). *Why is Mobility in India So Low? Social Insurance, Inequality, and Growth* (Working paper). National Bureau of Economic Research.

Munshi, K., & Rosenzweig, M. (2016). Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap. *The American Economic Review*, *106*(1), 46–98.

Myerson, R. B. (1983). *Analysis of Two Bargaining Problems with Incomplete Information.* Center for Mathematical Studies in Economics and Management Science, Northwestern University.

Myerson, R. B. (1984). Two-person Bargaining Problems with Incomplete Information. *Econometrica*, 461–487.

Myerson, R. B. (2005). Cooperative Games with Incomplete Information. In *Differential Information Economies* (pp. 481–505). Springer.

Myerson, R. B. (2007). Virtual Utility and the Core for Games with Incomplete Information. *Journal of Economic Theory*, 260–285.

Nier, E. W. (2005). Bank Stability and Transparency. *Journal of Financial Stability*, 342–354.

Palanski, M. E., Kahai, S. S., & Yammarino, F. J. (2011). Team Virtues and Performance: An Examination of Transparency, Behavioral Integrity, and Trust. *Journal of Business Ethics*, 201–216.

Parkhurst, J. T., & Hopmeyer, A. (1998). Sociometric Popularity and Peer-Perceived Popularity: Two Distinct Dimensions of Peer Status. *The Journal of Early Adolescence*, *18*(2), 125–144.

Paulson, A. L., Townsend, R. M., & Karaivanov, A. (2006). Distinguishing Limited Liability from Moral Hazard in a Model of Entrepreneurship. *Journal of Political Economy*, *114*(1), 100–144.

Porta, R. L., Lopez-De-Silane, F., Shleifer, A., & Vishny, R. W. (1997). Trust in Large Organizations. *American Economic Review*, *87*(2), 333–338.

Prachi, J. (2016). *Imperfect Monitoring and Informal Risk Sharing: The Role of Social Ties* (Working paper). Princeton University.

Putnam, R. D. (1995). Bowling Alone: America's Declining Social Capital. *Journal of Democracy*, *6*(1), 65–78.

Putnam, R. D., Leonardi, R., & Nanetti, R. Y. (1993). *Making Democracy Work: Civic Traditions in Modern Italy.* Princeton University Press.

Quibria, M. G., & Rashid, S. (1984). The Puzzle of Sharecropping: A Survey of Theories. *World Development*, *12*(2), 103–114.

Rahman, D. M. (2005). *Team Formation and Organization* (Unpublished doctoral dissertation). University of California Los Angeles.

Rayo, L. (2007). Relational Incentives and Moral Hazard in Teams. *The Review of Economic Studies*, 937–963.

Ricker-Gilbert, J. (2014). Wage and Employment Effects of Malawi's Fertilizer Subsidy Program. *Agricultural Economics*, *45*(3), 337–353.

Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007). An Introduction to Exponential Random Graph (p^*) Models for Social Networks. *Social Networks*, *29*(2), 173–191.

Rogerson, W. P. (1985). The First-Order Approach to Principal-Agent Problems. *Econometrica*, *53*(6), 1357–1367.

Rosenzweig, M. R., & Binswanger, H. P. (1993). Wealth, Weather Risk, and the Composition and Profitability of Agricultural Investments. *The Economic Journal*, *103*(416), 56–78.

Santos, P., & Barrett, C. B. (2010). Identity, Interest and Information Search in a Dynamic Rural Economy. *World Development*, *38*(12), 1788–1796.

Santos, P., & Barrett, C. B. (2011). Persistent Poverty and Informal Credit. *Journal of Development Economics*, *96*(2), 337–347.

Schlee, E. E. (2001). The Value of Information in Efficient Risk-Sharing Arrangements. *American Economic Review*, 509–524.

Sen, D. (2016). Sharecropping in Theory and Practice: A Selective Review. In *Understanding development* (pp. 53–72). Springer.

Singh, N. (1991). Theories of Sharecropping. In *The Economic Theory of Agrarian Institutions* (pp. 19–71). Oxford University Press.

Singh, R. P., Binswanger, H. P., & Jodha, N. S. (1985). *Manual of Instructions for Economic Investigators in ICRISAT's Village Level Studies* (Technical report). International Crops Research Institute for the Semi-Arid Tropics.

Slavin, R. E. (2011). Cooperative Learning. *Learning and Cognition in Education*, 160–166.

Sunding, D., & Zilberman, D. (2001). The Agricultural Innovation Process: Research and Technology Adoption in a Changing Agricultural Sector. In *Handbook of agricultural economics* (pp. 207–261). Elsevier.

Suri, T. (2011). *Estimating the Extent of Local Risk Sharing between Households* (Working paper). MIT Sloan School of Management.

Tadesse, S. (2006). The Economic Value of Regulated Disclosure: Evidence from the Banking Sector. *Journal of Accounting and Public Policy*, 32–70.

Townsend, R. M. (1994). Risk and Insurance in Village India. *Econometrica*, *62*(3), 539–591.

Tumennasan, N. (2011). *Do Contracts Help? A Team Formation Perspective* (Tech. Rep.).

Tumennasan, N. (2014). Moral Hazard and Stability. *Social Choice and Welfare*, *43*.

Udry, C. (2010). The Economics of Agriculture in Africa: Notes Toward a Research Program. *African Journal of Agricultural and Resource Economics*, *5*(1).

von Siemens, F., & Kosfeld, M. (2014). Team Production in Competitive Labor Markets with Adverse Selection. *European Economic Review*, 181–198.

Wang, X. Y. (2013). *A Note on Moral Hazard and Linear Compensation Schemes* (Working paper). Duke University.

Wimmer, A., & Lewis, K. (2010). Beyond and Below Racial Homophily: ERG Models of a Friendship Network Documented on Facebook. *American Journal of Sociology*, *116*(2), 583–642.

Woolcock, M. (1998). Social Capital and Economic Development: Toward a Theoretical Synthesis and Policy Framework. *Theory and Society*, *27*(2), 151–208.

Yamagishi, T., Cook, K. S., & Watabe, M. (1998). Uncertainty, Trust, and Commitment Formation in the United States and Japan. *American Journal of Sociology*, *104*(1), AJSv104p165–194.

Zak, P. J., & Knack, S. (2001). Trust and Growth. *The Economic Journal*, *111*(470), 295–321.

Tables and Figures: Chapter 1

Table A1: Summary Statistics

Variable	Average	Std. Dev.
Household size	4.97	2.13
Number of infants	0.05	0.23
Average adult age	40.76	8.97
Age-sex weight	4.31	1.69
Monthly consumption	117.87	73.60
Monthly income	115.29	147.51
Monthly effort	21.26	26.38
Monthly fertilizer	24.65	69.14
Number of households	876	
Observations	20044	

Notes: All money values in 1975 rupees. Consumption, income, effort, and fertilizer expressed in adult-equivalent terms. Household-month observations.

Table A2: A Test for Full Sharing

Dep. variable: $\log(c_{it})$	$\widehat{\beta}$ (s.e.)
$\log(y_{it})$	.2594*** (.0207)
Household fixed effects	Yes
Village-month fixed effects	Yes
R-squared	0.773
Observations	18,931

Notes: OLS regressions of log income on log consumption. Standard errors are robust.

Table A3: Risk Sharing and Effort

Dep. variable: $\log(c_{it})$	$\widehat{\beta}$ (s.e.)
$\log(y_{it})$	.1878*** (.0515)
$\log(y_{it}) \times \overline{\log(e_i)}$	.0278* (.0155)
Household fixed effects	Yes
Village-month fixed effects	Yes
R-squared	0.773
Observations	18,929

Notes: OLS regressions of log income and log income times average log effort on log consumption. Standard errors are robust.

Table A4: Non-Linearity between Risk Sharing and Effort

Dep. variable: $\log(c_{it})$	$\widehat{\beta}$ (s.e.)
$\log(y_{it})$	.1841*** (.0416)
$\log(y_{it}) \times \overline{e}_i$	.0048** (.0016)
$\log(y_{it}) \times \overline{e}_i^2$	-.0001** (.0001)
Household fixed effects	Yes
Village-month fixed effects	Yes
R-squared	0.774
Observations	18,931

Notes: OLS regressions of log income, log income times average effort, and log income times average effort squared on log consumption. Standard errors are robust.

Table A5: Effort and Fertilizer

Dep. variable: $\log(f_{it})$	$\widehat{\gamma}$ (s.e.)
$\log(e_{it})$	.4348*** (.0141)
Land area	.0032*** (.0003)
Household fixed effects	Yes
Village-month fixed effects	Yes
R-squared	0.326
Observations	8,794

Notes: OLS regressions of effort on fertilizer. All regressions with household and month fixed effects. Standard errors are robust.

Table A6: Risk Sharing and Fertilizer

Dep. variable: Consumption	$\widehat{\beta}$ (s.e.)
Income	.0575 (.1014)
Income × fertilizer	.0619** (.0269)
Household fixed effects	Yes
Village-month fixed effects	Yes
R-squared	0.774
Observations	18,890

Notes: OLS regressions of log income and log income times average log fertilizer on log consumption. Standard errors are robust.

Table A7: Non-Linearity between Risk Sharing and Fertilizer

Dep. variable: $\log(c_{it})$	$\widehat{\beta}$ (s.e.)
$\log(y_{it})$	.2143*** (.0416)
$\log(y_{it}) \times \overline{f}_i$	.0026 (.0022)
$\log(y_{it}) \times \overline{f}_i^2$	-.0001 (.0001)
Household fixed effects	Yes
Village-month fixed effects	Yes
R-squared	0.773
Observations	18,931

Notes: OLS regressions of log income, log income times average fertilizer, and log income times average fertilizer squared on log consumption. Standard errors are robust.

Table A8: Summary Statistics for $\widetilde{\log\left(\frac{f_{it}}{e_{it}}\right)}$.

	Average	S.d.	Min	Max
$\widetilde{\alpha}_{vt} = 0$	.3568	.9308	-3.0420	6.4878
$\widetilde{\alpha}_{vt} = 1$	1.7491	.9108	-1.5723	7.7160

Table A9: Summary Statistics of the Growth Rates of Effort and Fertilizer Use (from $\widetilde{\alpha}_{vt} = 0$ to $\widetilde{\alpha}_{vt} = 1$).

	Average	S.d.	Min	Max
$\log(e_{it}(0)) - \log(e_{it}(1))$	-3.4276	.5628	-4.4330	-1.4126
$\log(f_{it}(0)) - \log(f_{it}(1))$	-2.0353	.5212	-2.8505	-.0481

Table A10: Which Households Are More Affected by Risk Sharing?

Dep. variable: $\widetilde{x}_{it}(1) - \widetilde{x}_{it}(0)$	$\widehat{\beta}$ (s.e.)
Household size	.0188*** (.0043)
Plot area	.0011** (.0004)
y_{it}	.0001** (.0001)
Household fixed effects	Yes
Village-month fixed effects	Yes
R-squared	0.999
Observations	6864

Notes: OLS regressions of household size, average household age, plot area, and income on the growth rate of fertilizer used per hours worked when going from no sharing to full insurance. Standard errors are robust.

Table A11: A Test for Full Sharing with Simulated Consumption

Dep. variable:	$\log(\widetilde{c}_{it})$ $(\widetilde{\alpha}_{vt} = 0.82)$	$\widehat{\beta}$ (s.e.)	$\log(\widetilde{c}_{it})$ $(\widetilde{\alpha}_{vt} = 0.75)$	$\widehat{\beta}$ (s.e.)	$\log(\widetilde{c}_{it})$ $(\widehat{\alpha}_{vt})$	$\widehat{\beta}$ (s.e.)
$\log(y_{it})$		.1600*** (.0142)		.2123*** (.0187)		.2968*** (.0662)
Household fixed effects		Yes		Yes		Yes
Village-month fixed effects		Yes		Yes		Yes
R-squared		0.967		0.962		0.817
Observations		15,069		15,095		5,465

Notes: OLS regressions of log income on log of simulated consumption. Standard errors are robust.

Table A12: Within-Estimator

Dep. variable: $c_{it} - \overline{c}_{vt}$	$\widehat{\beta}$ (s.e.)
$\pi_{it} - \overline{\pi}_{vt}$	.1731*** (.0029)
Household fixed effects	No
Village-month fixed effects	No
R-squared	0.1519
Observations	20044

Notes: OLS regressions of deviations of household income from village-month average income on deviations household consumption from village-month average consumption.

Table A13: Between-Estimator

Dep. variable: $c_{it} - \overline{c}_{vt}$	$\widehat{\beta}$ (s.e.)
$\pi_{it} - \overline{\pi}_{vt}$	.9228*** (.0021)
Household fixed effects	No
Village-month fixed effects	No
R-squared	0.8978
Observations	21080

Notes: OLS regressions of village-month average income on village-month average consumption.

Figure 1.1: Histogram of $\log\left(\widehat{k}_i\right)$

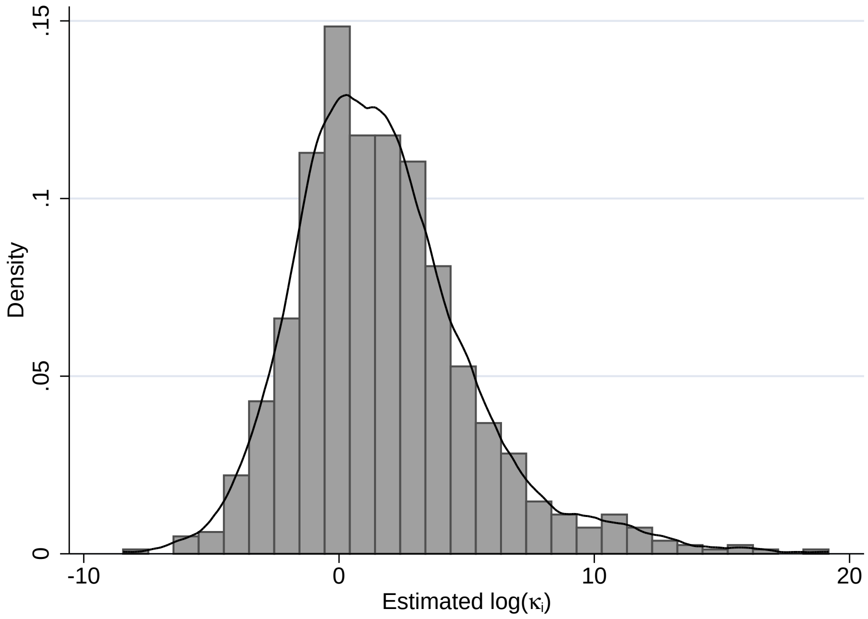


Figure 1.2: Histogram of $\widehat{\alpha}_{vt}$

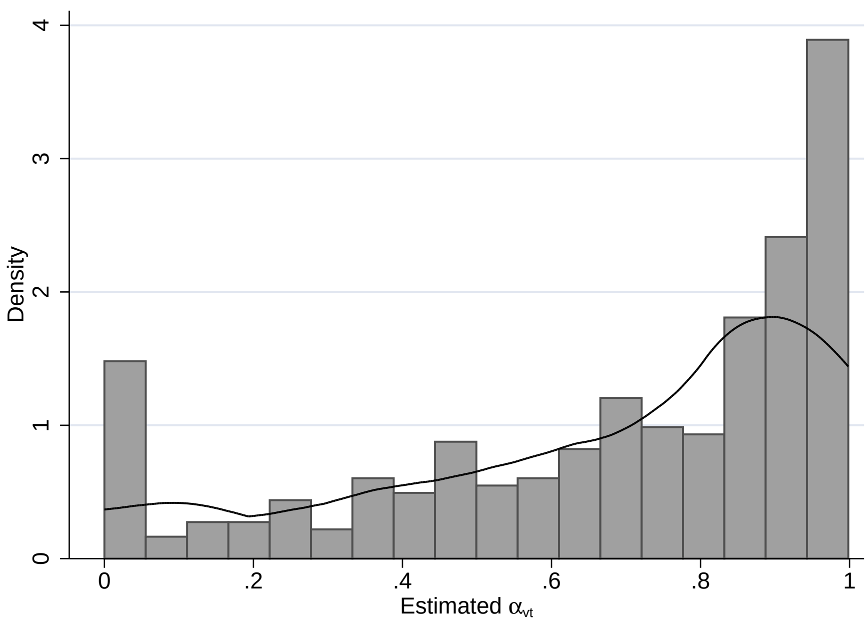


Figure 1.3: Comparative Statics

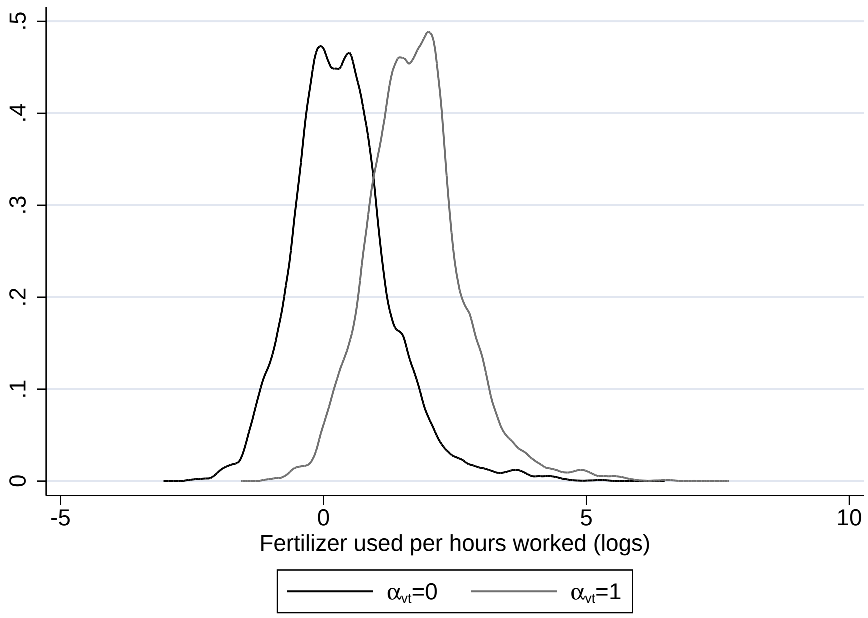


Figure 1.4: A Validation Exercise

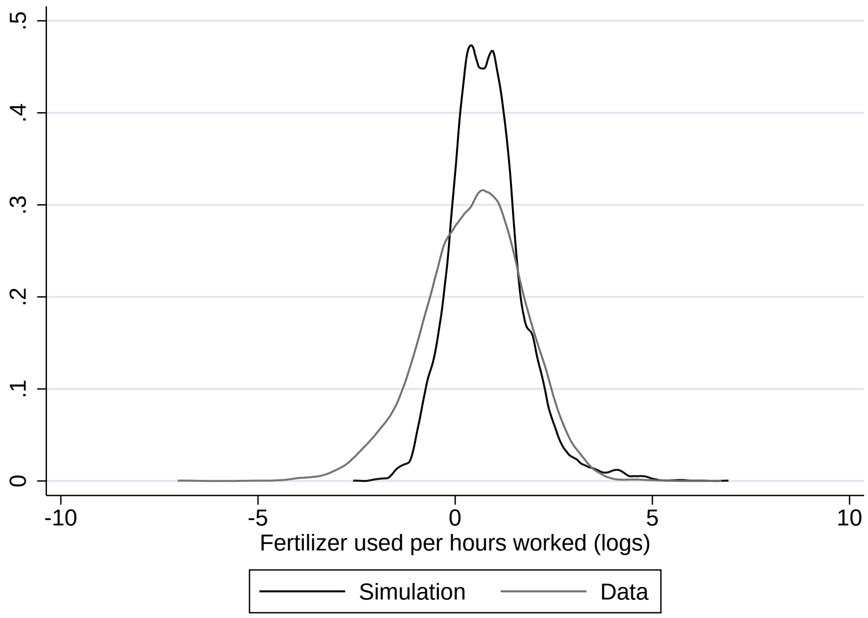


Figure 1.5: Optimal Sharing

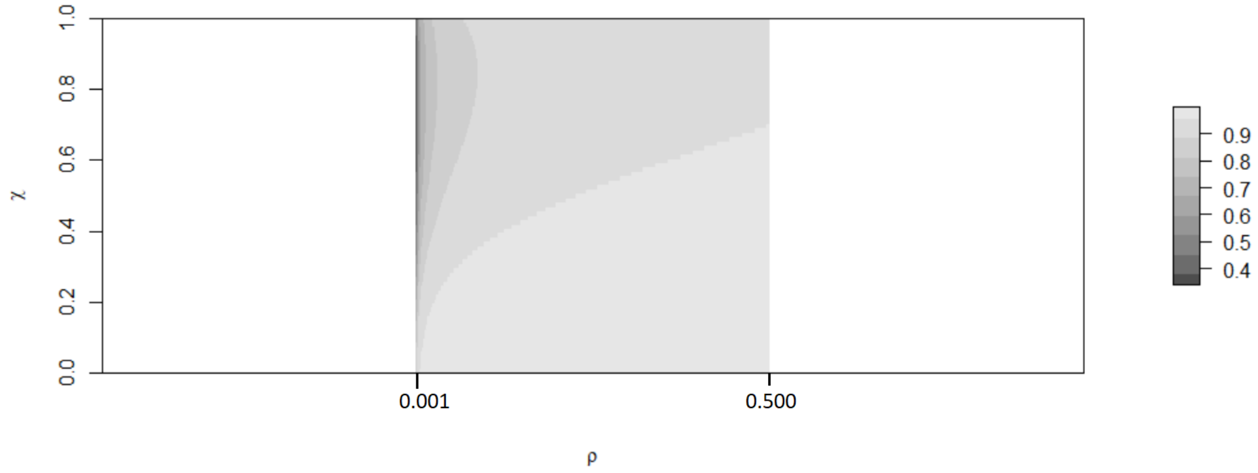
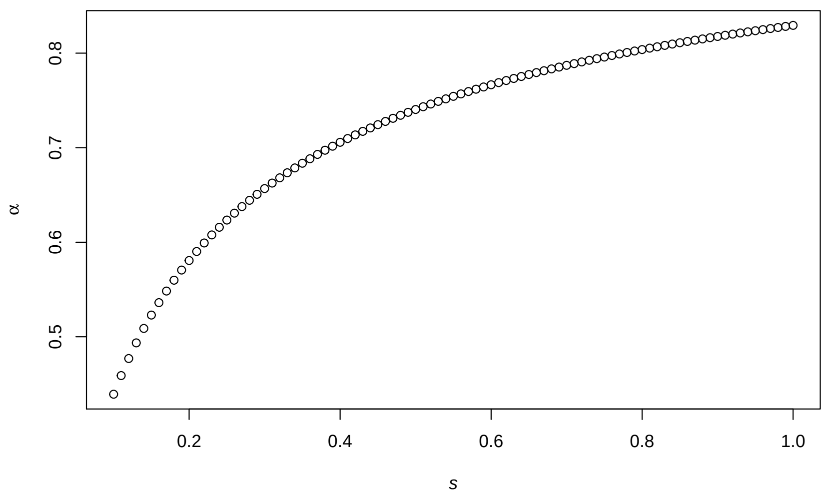


Figure 1.6: Optimal Sharing and Fertilizer Subsidy



Tables and Figures: Chapter 2

Table B.1: Summary Statistics for the Entire Population

Variable	Avg.	S.d.	Min	Max	N
Siblings	1.32	.83	0	4	100
Friends at Uniandes met before starting university	10.66	10.02	0	50	100
Friends at Uniandes met after starting university	9.89	8.35	0	40	100
Friends not at Uniandes	16.59	17.31	0	100	99
Hr/wk spent socializing with friends at Uniandes	10.40	9.14	0	65	100
Hr/wk spent socializing with friends not at Uniandes	7.25	11.38	0	85	100
Hr/wk of physical activity	4.39	4.27	0	21	100
Realistic	3.17	1.18	1	5	100
Introverted	2.93	1.07	1	5	100
Inhibited	2.49	.95	1	5	100
Shy	2.33	1.08	1	5	100
Age	18.40	1.07	16	26	96
Height (cm)	172.14	14.47	60	192	96
Weight (kg)	65.4	9.90	46	100	95
Glasses (1 = yes)	.30	.46	0	1	109
Tattoos (1 = yes)	.04	.19	0	1	109
Piercings (1 = yes)	.19	.40	0	1	109
Smoker (1 = yes)	.32	.47	0	1	109
Go to parties (1 = yes)	.85	.36	0	1	109

Table B.2: Summary Statistics for the Social Networks

Network	Variable	Avg.	S.d.	Min	Max
Saying hello	Degree (In)	13.5	6.29	1	27
	Degree (Out)	13.15	7.19	1	38
	Clustering (Avg.)	0.52	0.12	0	0.93
	Clustering (Global)	0.49			
	Support	0.99	0.11	0	1
	Path length	1.7	0.5	1	3
Friendship	Degree (In)	6.85	3.60	0	14
	Degree (Out)	6.85	4.81	0	19
	Clustering (Avg.)	0.44	0.21	0	1
	Clustering (Global)	0.38			
	Support	0.96	0.2	0	1
	Path length	2.16	0.75	1	5
Studying together	Degree (In)	5.77	3.18	0	13
	Degree (Out)	5.77	4.91	0	29
	Clustering (Avg.)	0.44	0.2	0	1
	Clustering (Global)	0.33			
	Support	0.94	0.23	0	1
	Path length	2.13	0.69	1	4
Lunch together	Degree (In)	2.58	2.05	0	7
	Degree (Out)	2.58	2.32	0	11
	Clustering (Avg.)	0.4	0.3	0	1
	Clustering (Global)	0.38			
	Support	0.76	0.43	0	1
	Path length	4	2.04	1	12
Confiding	Degree (In)	2.98	2.12	0	7
	Degree (Out)	2.98	2.53	0	10
	Clustering (Avg.)	0.35	0.28	0	1
	Clustering (Global)	0.31			
	Support	0.76	0.43	0	1
	Path length	2.89	1.09	1	6
Know before	Degree (In)	1.49	2.15	0	7
	Degree (Out)	1.49	2.28	0	8
	Clustering (Avg.)	0.87	0.27	0	1
	Clustering (Global)	0.85			
	Support	0.72	0.45	0	1
	Path length	1.31	0.5	1	3

Table B.3: "Friendship" Network on Trust.

	Coefficient	Marginal effect	QAP p-value
Constant	20.66		0.37
Trust i	0.05	0.0046	0.56
Trust j	0.051	0.0047	0.29
Trust Δ	0.03	0	0.63
Controls	Yes		
AIC	1868.45		
BIC	2318.48		
Deviance	1716.45		
Null Deviance	3820.63		
Observations	2756		

* indicates significance at the 5% level. ** indicates significance at the 1% level.

Table B.4: "Greeting Each Other" Network on Trust.

	Coefficient	Marginal effect	QAP p-value
Constant	13.86		0.39
Trust i	0.04	0.01	0.65
Trust j	0.02	0	0.75
Trust Δ	0.04	0.01	0.34
Controls	Yes		
AIC	2495.23		
BIC	2945.26		
Deviance	2343.23		
Null Deviance	3820.63		
Observations	2756		

* indicates significance at the 5% level. ** indicates significance at the 1% level.

Table B.5: "Studying Together" Network on Trust.

	Coefficient	Marginal effect	QAP p-value
Constant	18.17		0.42
Trust i	-0.02	0	0.78
Trust j	0	0	0.98
Trust Δ	0.02	0	0.74
Controls	Yes		
AIC	1642.32		
BIC	2092.36		
Deviance	1490.32		
Null Deviance	3820.63		
Observations	2756		

* indicates significance at the 5% level. ** indicates significance at the 1% level.

Table B.6: "Confiding in" Network on Trust

	Coefficient	Marginal effect	QAP p-value
Constant	27.33		0.29
Trust i	0	0	0.99
Trust j	0.1	0	0.12
Trust Δ	0.07	0	0.35
Controls	Yes		
AIC	1044.64		
BIC	1494.67		
Deviance	892.64		
Null Deviance	3820.63		
Observations	2756		

* indicates significance at the 5% level. ** indicates significance at the 1% level.

Table B.7: "Having Lunch Together" Network on Trust.

	Coefficient	Marginal effect	QAP p-value
Constant	52.11		0.09
Trust i	-0.03	0	0.8
Trust j	-0.01	0	0.89
Trust Δ	-0.01	0	0.92
Controls	Yes		
AIC	967.75		
BIC	1417.78		
Deviance	815.75		
Null Deviance	3820.63		
Observations	2756		

* indicates significance at the 5% level. ** indicates significance at the 1% level.

Table B.8: "Friendship" Network on Trust, All Controls Displayed

	Coefficient	Marginal effect	QAP p-value
Constant	20.66		0.37
Knew each other from before	2.86**	0.26**	0
Time of exposure	0.13**	0.01**	0
Trust i	0.05	0	0.56
Trust j	0.05	0	0.29
Trust Δ	0.03	0	0.63
Socio-economic background Δ	−0.31**	−0.03**	0
Socio-economic background i	−0.18	−0.02	0.43
Socio-economic background j	−0.19	−0.02	0.33
Sex i	−0.42	−0.04	0.52
Sex j	−0.33	−0.03	0.45
Sex Δ	−0.04	0	0.85
Friends met before i	−0.01	0	0.5
Friends met before j	−0.01	0	0.37
Friends met before Δ	−0.02	0	0.08
GPA i	0.66	0.06	0.35
GPA j	0.64	0.06	0.15
GPA Δ	−0.5	−0.05	0.2
High school exams i	−0.01	0	0.46
High school exams j	−0.02	0	0.1
High school exams Δ	−0.01	0	0.43
Age i	−0.5	−0.05	0.32
Age j	−0.31	−0.03	0.42
Age Δ	0.06	0.01	0.77
Weight i	−0.04	0	0.16
Weight j	−0.02	0	0.48
Weight Δ	−0.02	0	0.39
Glasses i	0.03	0	0.9
Glasses j	−0.52	−0.05	0.17
Glasses Δ	−0.01	0	0.92
Brown eyes i	−0.15	−0.01	0.84
Blue eyes i	−0.33	−0.03	0.76
Green eyes i	−0.02	0	0.96
Other color eyes i	0.7	0.06	0.58
Brown eyes j	0.29	0.03	0.57

Blue eyes j	0.45	0.04	0.49
Green eyes j	0.5	0.05	0.46
Other color eyes j	1.38	0.13	0.18
Eyes Different	0.18	0.02	0.43
Brown hair i	-1.17^{*}	-0.11^{*}	0.03
Blonde hair i	−1.9	−0.17	0.23
Other color hair i	−2.6	−0.24	0.12
Brown hair j	−0.7	−0.06	0.08
Blonde hair j	−1.83	−0.17	0.05
Other color hair j	−0.96	−0.09	0.33
Hair Different	−0.26	−0.02	0.15
Height i	0.05	0	0.32
Height j	0	0	0.97
Height Δ	−0.02	0	0.31
Residence i	−0.6	−0.06	0.28
Residence j	−0.7	−0.06	0.05
Residence Δ	-0.54^{*}	-0.05^{*}	0.01
Piercings i	−0.86	−0.08	0.16
Piercings j	−0.3	−0.03	0.56
Piercings Δ	−0.32	−0.03	0.19
Attending parties i	−0.44	−0.04	0.37
Attending parties j	0.26	0.02	0.45
Attending parties Δ	0.07	0.01	0.65
Siblings i	−0.17	−0.02	0.53
Siblings j	−0.08	−0.01	0.64
Siblings Δ	0.08	0.01	0.5
Realistic personality i	−0.04	0	0.78
Realistic personality j	0.16	0.01	0.33
Realistic personality Δ	0.06	0.01	0.53
Introvert personality i	−0.61	−0.06	0.06
Introvert personality j	-0.4^{**}	-0.04^{**}	0
Introvert personality Δ	0.08	0.01	0.59
Inhibited personality i	0.06	0.01	0.79
Inhibited personality j	0	0	0.99
Inhibited personality Δ	−0.19	−0.02	0.05
Shy personality i	−0.32	−0.03	0.31
Shy personality j	0.14	0.01	0.53

Shy personality Δ	0.11	0.01	0.43
Physical activity i	−0.05	0	0.31
Physical activity j	−0.06	−0.01	0.14
Physical activity Δ	0.06	0.01	0.1
AIC	1868.45		
BIC	2318.48		
Deviance	1716.45		
Null Deviance	3820.63		
Observations	2756		

* indicates significance at the 5% level. ** indicates significance at the 1% level.

Table B.9: "Studying Together" Network on Trust, All Controls Displayed

	Coefficient	Marginal effect	QAP p-value
Constant	18.17		0.42
Knew each other from before	2.17**	0.17**	0
Time of exposure	0.18**	0.01**	0
Trust i	−0.02	0	0.78
Trust j	0	0	0.98
Trust Δ	0.02	0	0.74
Socio-economic background Δ	−0.56**	−0.04**	0
Socio-economic background i	−0.06	0	0.83
Socio-economic background j	−0.27	−0.02	0.13
Sex i	−0.3	−0.02	0.69
Sex j	−0.44	−0.03	0.36
Sex Δ	−0.29	−0.02	0.2
Friends met before i	−0.03	0	0.08
Friends met before j	0	0	0.76
Friends met before Δ	−0.02*	0*	0.03
GPA i	−0.54	−0.04	0.55
GPA j	0.75	0.06	0.24
GPA Δ	−0.62	−0.05	0.12
High school exams i	0	0	0.85
High school exams j	−0.01	0	0.25
High school exams Δ	0	0	0.89
Age i	−0.67	−0.05	0.28
Age j	−0.41	−0.03	0.25
Age Δ	−0.33	−0.03	0.22
Weight i	0.01	0	0.78
Weight j	−0.04	0	0.09
Weight Δ	−0.01	0	0.64
Glasses i	0.08	0.01	0.89
Glasses j	−0.42	−0.03	0.33
Glasses Δ	−0.02	0	0.97
Brown eyes i	0.19	0.02	0.84
Blue eyes i	−0.39	−0.03	0.8
Green eyes i	−0.31	−0.02	0.74
Other color eyes i	0.17	0.01	0.93
Brown eyes j	0.79	0.06	0.13

Blue eyes j	0.89	0.07	0.37
Green eyes j	0.26	0.02	0.79
Other color eyes j	1.35	0.11	0.16
Eyes Different	0.15	0.01	0.53
Brown hair i	−0.12	−0.01	0.84
Blonde hair i	−1.21	−0.1	0.49
Other color hair i	1.67	0.13	0.46
Brown hair j	−0.48	−0.04	0.17
Blonde hair j	−1.98	−0.16	0.09
Other color hair j	−0.42	−0.03	0.72
Hair Different	−0.33	−0.03	0.07
Height i	0.02	0	0.8
Height j	0.04	0	0.27
Height Δ	−0.02	0	0.41
Residence i	0.62	0.05	0.37
Residence j	−0.57	−0.04	0.16
Residence Δ	−0.29	−0.02	0.21
Piercings i	−0.02	0	0.98
Piercings j	−0.43	−0.03	0.32
Piercings Δ	−0.46	−0.04	0.05
Attending parties i	0.53	0.04	0.48
Attending parties j	0.23	0.02	0.53
Attending parties Δ	−0.1	−0.01	0.62
Siblings i	−0.42	−0.03	0.25
Siblings j	−0.27	−0.02	0.25
Siblings Δ	0.23	0.02	0.1
Realistic personality i	0.04	0	0.84
Realistic personality j	0.15	0.01	0.25
Realistic personality Δ	0.26*	0.02*	0.02
Introvert personality i	−0.62	−0.05	0.06
Introvert personality j	−0.46*	−0.04*	0.02
Introvert personality Δ	0	0	0.96
Inhibited personality i	−0.05	0	0.93
Inhibited personality j	0.01	0	0.97
Inhibited personality Δ	−0.1	−0.01	0.39
Shy personality i	0.32	0.03	0.49
Shy personality j	0.1	0.01	0.72

Shy personality Δ	0.14	0.01	0.25
Physical activity i	−0.06	0	0.42
Physical activity j	−0.07	−0.01	0.14
Physical activity Δ	0.05	0	0.2
AIC	1642.32		
BIC	2092.36		
Deviance	1490.32		
Null Deviance	3820.63		
Observations	2756		

* indicates significance at the 5% level. ** indicates significance at the 1% level.

Table B.10: "Confiding in" Network on Trust, All Controls Displayed

	Coefficient	Marginal effect	QAP p-value
Constant	27.33		0.29
Knew each other from before	2.78**	0.12**	0
Time of exposure	0.17**	0.01**	0
Trust i	0	0	0.99
Trust j	0.1	0	0.12
Trust Δ	0.07	0	0.35
Socio-economic background Δ	−0.44**	−0.02**	0
Socio-economic background i	−0.27	−0.01	0.36
Socio-economic background j	−0.36	−0.02	0.05
Sex i	−1.18	−0.05	0.14
Sex j	−1.25*	−0.05*	0.02
Sex Δ	−0.83*	−0.04*	0.02
Friends met before i	−0.01	0	0.64
Friends met before j	−0.01	0	0.52
Friends met before Δ	−0.01	0	0.35
GPA i	−0.86	−0.04	0.31
GPA j	0.31	0.01	0.64
GPA Δ	−1.08	−0.05	0.05
High school exams i	0.01	0	0.45
High school exams j	−0.01	0	0.36
High school exams Δ	0	0	0.94
Age i	−0.95	−0.04	0.13
Age j	−0.43	−0.02	0.36
Age Δ	0.18	0.01	0.63
Weight i	−0.01	0	0.87
Weight j	−0.02	0	0.44
Weight Δ	−0.02	0	0.32
Glasses i	−0.99	−0.04	0.15
Glasses j	−0.46	−0.02	0.38
Glasses Δ	−0.27	−0.01	0.36
Brown eyes i	0.41	0.02	0.64
Blue eyes i	0.7	0.03	0.6
Green eyes i	0.2	0.01	0.86
Other color eyes i	−2.01	−0.09	0.29
Brown eyes j	0.98	0.04	0.15

Blue eyes j	0.8	0.03	0.38
Green eyes j	1.05	0.05	0.27
Other color eyes j	0.88	0.04	0.52
Eyes Different	-0.2	-0.01	0.61
Brown hair i	-0.61	-0.03	0.24
Blonde hair i	-3.68	-0.16	0.06
Other color hair i	-2.32	-0.1	0.28
Brown hair j	-1.18^{**}	-0.05^{**}	0
Blonde hair j	-2.47	-0.11	0.05
Other color hair j	-2.27	-0.1	0.1
Hair Different	-0.27	-0.01	0.28
Height i	0.03	0	0.69
Height j	0.01	0	0.69
Height Δ	0.01	0	0.79
Residence i	-0.19	-0.01	0.79
Residence j	-1.29^{*}	-0.06^{*}	0.01
Residence Δ	-0.55	-0.02	0.05
Piercings i	-0.51	-0.02	0.51
Piercings j	-0.8	-0.03	0.23
Piercings Δ	-0.33	-0.01	0.35
Attending parties i	0.42	0.02	0.44
Attending parties j	-0.07	0	0.84
Attending parties Δ	-0.08	0	0.69
Siblings i	-0.12	-0.01	0.71
Siblings j	-0.13	-0.01	0.54
Siblings Δ	0.12	0	0.62
Realistic personality i	0.05	0	0.79
Realistic personality j	0.09	0	0.63
Realistic personality Δ	0.2	0.01	0.09
Introvert personality i	-0.82^{*}	-0.04^{*}	0.02
Introvert personality j	-0.47^{**}	-0.02^{**}	0
Introvert personality Δ	-0.03	0	0.85
Inhibited personality i	-0.1	0	0.83
Inhibited personality j	-0.13	-0.01	0.57
Inhibited personality Δ	-0.26	-0.01	0.11
Shy personality i	0.48	0.02	0.18
Shy personality j	0.08	0	0.81

Shy personality Δ	0.36	0.02	0.05
Physical activity i	−0.07	0	0.33
Physical activity j	−0.08	0	0.22
Physical activity Δ	0.09*	0*	0.04
AIC	1044.64		
BIC	1494.67		
Deviance	892.64		
Null Deviance	3820.63		
Observations	2756		

* indicates significance at the 5% level. ** indicates significance at the 1% level.

Table B.11: "Having Lunch Together" Network on Trust, All Controls Displayed

	Coefficient	Marginal effect	QAP p-value
Constant	52.11		0.09
Knew each other from before	2.05**	0.08**	0
Time of exposure	0.2**	0.01**	0
Trust i	−0.03	0	0.8
Trust j	−0.01	0	0.89
Trust Δ	−0.01	0	0.92
Socio-economic background Δ	−0.35*	−0.01*	0.03
Socio-economic background i	−0.31	−0.01	0.29
Socio-economic background j	−0.32	−0.01	0.12
Sex i	−0.15	−0.01	0.92
Sex j	−1.01	−0.04	0.05
Sex Δ	−0.46	−0.02	0.19
Friends met before i	−0.01	0	0.55
Friends met before j	0	0	0.87
Friends met before Δ	−0.01	0	0.53
GPA i	−0.58	−0.02	0.53
GPA j	−0.31	−0.01	0.6
GPA Δ	−0.99	−0.04	0.11
High school exams i	0	0	0.84
High school exams j	−0.01	0	0.38
High school exams Δ	0	0	0.96
Age i	−0.9	−0.03	0.16
Age j	−0.74	−0.03	0.22
Age Δ	0	0	1
Weight i	−0.01	0	0.86
Weight j	0.01	0	0.58
Weight Δ	−0.01	0	0.75
Glasses i	−0.35	−0.01	0.56
Glasses j	−0.58	−0.02	0.26
Glasses Δ	−0.14	−0.01	0.64
Brown eyes i	0.18	0.01	0.86
Blue eyes i	0.56	0.02	0.63
Green eyes i	0.32	0.01	0.71
Other color eyes i	−0.64	−0.02	0.75
Brown eyes j	0.18	0.01	0.81

Blue eyes j	0.83	0.03	0.38
Green eyes j	0.35	0.01	0.67
Other color eyes j	1.02	0.04	0.49
Eyes Different	0.09	0	0.84
Brown hair i	−0.41	−0.02	0.44
Blonde hair i	−2.91	−0.11	0.18
Other color hair i	−0.72	−0.03	0.7
Brown hair j	−0.99	−0.04	0.05
Blonde hair j	−2.57	−0.1	0.05
Other color hair j	−1.21	−0.05	0.46
Hair Different	−0.64	−0.02	0.05
Height i	−0.02	0	0.74
Height j	−0.04	0	0.4
Height Δ	−0.01	0	0.72
Residence i	−0.13	−0.01	0.85
Residence j	−0.78	−0.03	0.17
Residence Δ	−0.49	−0.02	0.13
Piercings i	0.1	0	0.91
Piercings j	−0.57	−0.02	0.33
Piercings Δ	−0.8	−0.03	0.05
Attending parties i	0.57	0.02	0.47
Attending parties j	0.33	0.01	0.49
Attending parties Δ	−0.1	0	0.73
Siblings i	0	0	1
Siblings j	−0.16	−0.01	0.56
Siblings Δ	0.06	0	0.8
Realistic personality i	0.02	0	0.93
Realistic personality j	0.16	0.01	0.4
Realistic personality Δ	0.2	0.01	0.24
Introvert personality i	-0.76^{*}	-0.03^{*}	0.02
Introvert personality j	-0.43^{**}	-0.02^{**}	0
Introvert personality Δ	0.05	0	0.82
Inhibited personality i	−0.1	0	0.84
Inhibited personality j	−0.25	−0.01	0.28
Inhibited personality Δ	−0.26	−0.01	0.12
Shy personality i	0.47	0.02	0.28
Shy personality j	0.3	0.01	0.32

Shy personality Δ	0.23	0.01	0.22
Physical activity i	−0.11	0	0.15
Physical activity j	−0.14*	−0.01*	0.04
Physical activity Δ	0.1*	0*	0.04
AIC	967.75		
BIC	1417.78		
Deviance	815.75		
Null Deviance	3820.63		
Observations	2756		

* indicates significance at the 5% level. ** indicates significance at the 1% level.

Table B.12: "Greeting Each Other" Network on Trust, All Controls Displayed

	Coefficient	Marginal effect	QAP p-value
Constant	13.86		0.39
Knew each other from before	4.91**	0.67**	0
Time of exposure	0.17**	0.02**	0
Trust i	0.04	0.01	0.65
Trust j	0.02	0	0.75
Trust Δ	0.04	0.01	0.34
Socio-economic background Δ	−0.35**	−0.05**	0
Socio-economic background i	−0.15	−0.02	0.51
Socio-economic background j	−0.19	−0.03	0.24
Sex i	−0.62	−0.08	0.38
Sex j	−0.59	−0.08	0.21
Sex Δ	−0.12	−0.02	0.41
Friends met before i	0	0	0.97
Friends met before j	0	0	0.78
Friends met before Δ	−0.02**	0**	0
GPA i	0.55	0.07	0.35
GPA j	0.69	0.09	0.07
GPA Δ	−0.59	−0.08	0.08
High school exams i	−0.01	0	0.35
High school exams j	−0.01	0	0.13
High school exams Δ	0	0	0.7
Age i	−0.61	−0.08	0.21
Age j	−0.04	−0.01	0.92
Age Δ	−0.04	0	0.86
Weight i	−0.04	−0.01	0.23
Weight j	−0.02	0	0.35
Weight Δ	−0.01	0	0.47
Glasses i	0.03	0	0.95
Glasses j	−0.6	−0.08	0.07
Glasses Δ	0.03	0	0.85
Brown eyes i	−0.4	−0.06	0.56
Blue eyes i	−0.07	−0.01	0.96
Green eyes i	−0.36	−0.05	0.74
Other color eyes i	0.11	0.02	0.93
Brown eyes j	0.57	0.08	0.28

Blue eyes j	0.46	0.06	0.46
Green eyes j	0.56	0.08	0.52
Other color eyes j	1.75	0.24	0.08
Eyes Different	0.05	0.01	0.72
Brown hair i	-0.57	-0.08	0.26
Blonde hair i	-1.89	-0.26	0.15
Other color hair i	-1.86	-0.25	0.29
Brown hair j	-0.76^{*}	-0.1^{*}	0.03
Blonde hair j	-2.19	-0.3	0.06
Other color hair j	-0.21	-0.03	0.95
Hair Different	-0.32^{*}	-0.04^{*}	0.01
Height i	0.05	0.01	0.38
Height j	0.01	0	0.76
Height Δ	-0.03	0	0.05
Residence i	-0.21	-0.03	0.66
Residence j	-1.03^{*}	-0.14^{*}	0.01
Residence Δ	-0.34^{*}	-0.05^{*}	0.01
Piercings i	-0.82	-0.11	0.17
Piercings j	-0.21	-0.03	0.71
Piercings Δ	-0.12	-0.02	0.48
Attending parties i	0.21	0.03	0.66
Attending parties j	0.14	0.02	0.77
Attending parties Δ	-0.12	-0.02	0.43
Siblings i	0.02	0	0.97
Siblings j	-0.13	-0.02	0.51
Siblings Δ	0.01	0	0.94
Realistic personality i	0.14	0.02	0.45
Realistic personality j	0.16	0.02	0.34
Realistic personality Δ	0.02	0	0.84
Introvert personality i	-0.69^{**}	-0.09^{**}	0
Introvert personality j	-0.51^{**}	-0.07^{**}	0
Introvert personality Δ	0.03	0	0.79
Inhibited personality i	0.19	0.03	0.48
Inhibited personality j	0.02	0	0.92
Inhibited personality Δ	-0.21	-0.03	0.09
Shy personality i	-0.18	-0.02	0.51
Shy personality j	0.18	0.02	0.38

Shy personality Δ	0.01	0	0.94
Physical activity i	−0.03	0	0.64
Physical activity j	−0.06	−0.01	0.13
Physical activity Δ	0.03	0	0.36
AIC	2495.23		
BIC	2945.26		
Deviance	2343.23		
Null Deviance	3820.63		
Observations	2756		

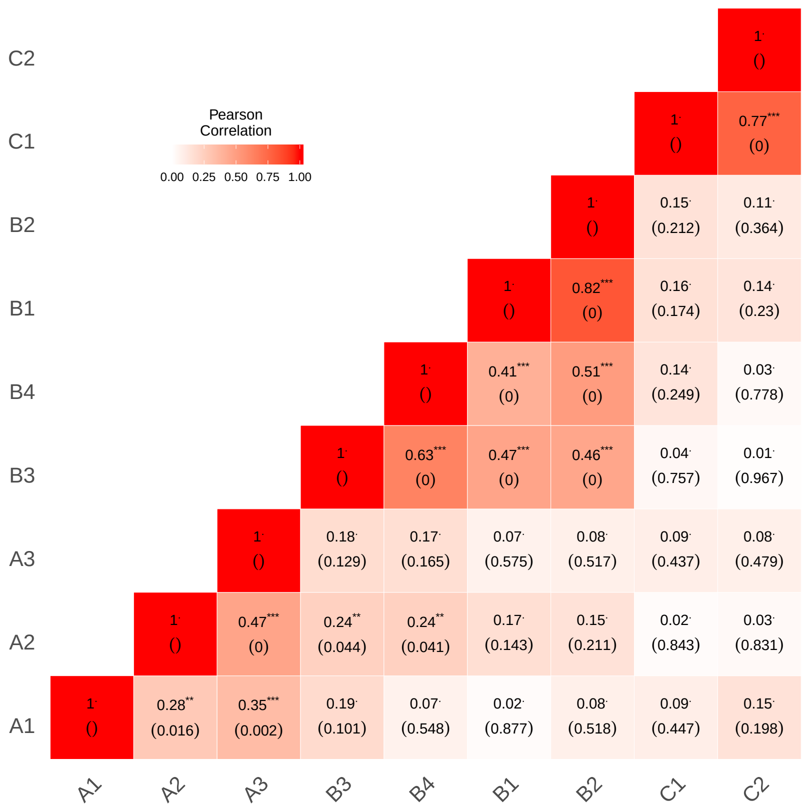
* indicates significance at the 5% level. ** indicates significance at the 1% level.

Table B.13: ERGM: "Friendship" Networks on Several Trust Measures.

	Directed	Union	Intersection
Constant	-6.682^{***}	-6.530^{***}	-5.743^{***}
	(0.072)	(0.065)	(0.070)
Generalized trust	-0.026	-0.031	-0.054^{*}
	(0.018)	(0.024)	(0.032)
Trust towards friends	0.005	0.009	0.003
	(0.006)	(0.010)	(0.014)
Trust towards neighbors	-0.010	-0.012	-0.022
	(0.007)	(0.013)	(0.019)
Socio-economic background Δ	-0.352^{***}	-0.399^{**}	-0.748^{***}
	(0.106)	(0.172)	(0.241)
Knew each other from before	1.766^{***}	2.462^{***}	2.040^{***}
	(0.199)	(0.370)	(0.320)
Controls	Yes	Yes	Yes
Observations	2756	2756	2756

* indicates significance at the 5% level. ** indicates significance at the 1% level.

Figure 2.1: Correlations Among Answers to Questions $A1$–$C2$.



p values displayed in parentheses.
Correlation coefficients with two asterisk indicate statistically significant relationships at alpha=0.05, and three asterisks mean significant at alpha=0.01

Figure 2.2: "Friendship" and "Greeting Each Other" Networks.

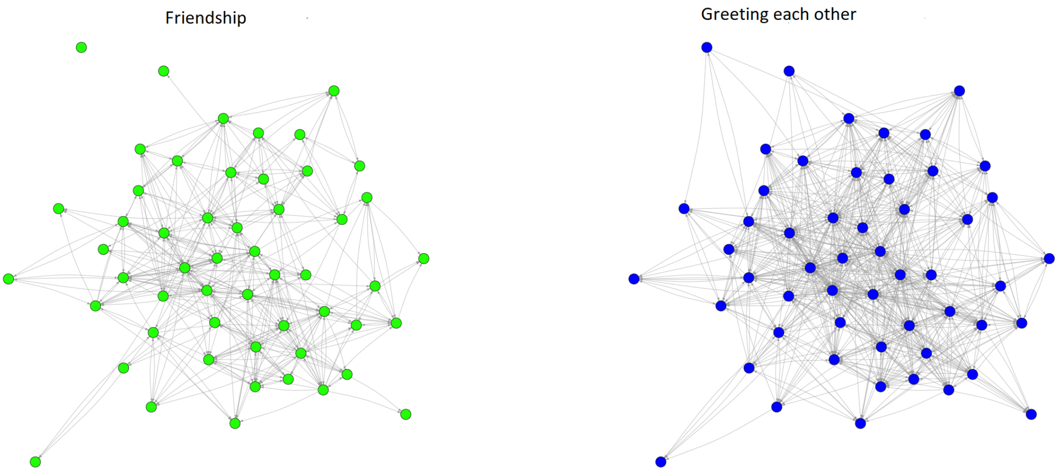


Figure 2.3: "Having Lunch Together" and "Confiding in" Networks.

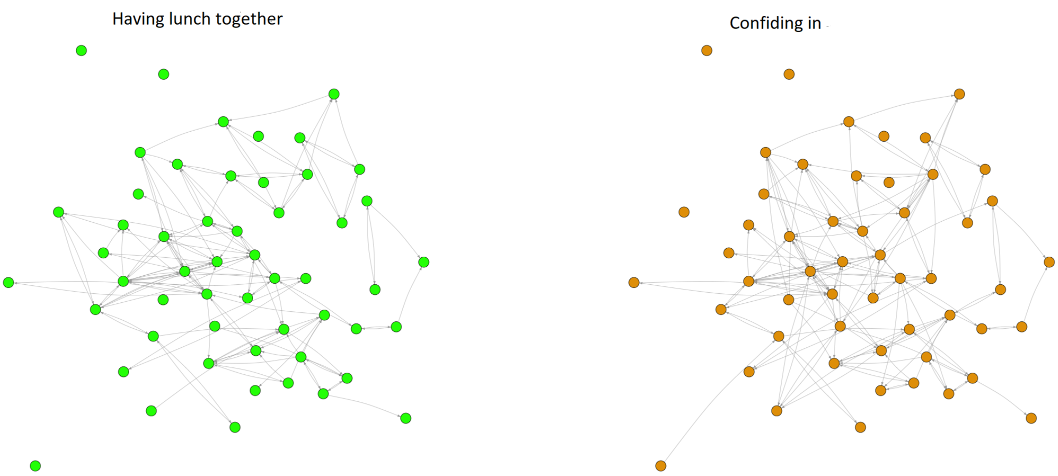
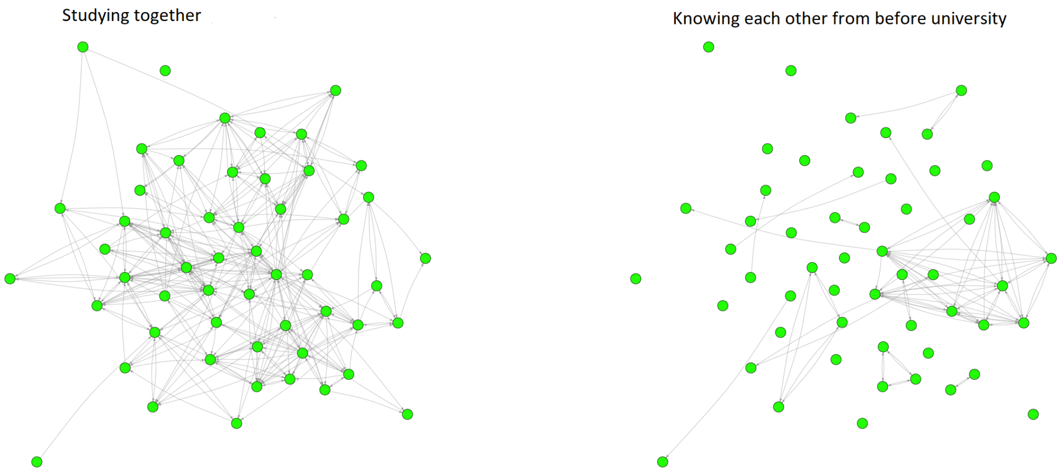


Figure 2.4: “Studying Together” and “Knowing Each Other from Before University” Networks.



Appendix: Chapter 1

A.1 A Full-Fledged Model of Risk Sharing

In this appendix, I extend the model outlined in Section 1.2 by dropping the assumption that sharing contracts are linear. Moreover, instead of focusing solely on fertilizer use, I assume that each household chooses the quantities of m different agricultural inputs (intermediates) $\boldsymbol{z}_i \in \mathbb{R}_+^m$, whose prices $\boldsymbol{p} \in \mathbb{R}_{++}^m$ the household takes as given.

As before, consider n household-farms, each of which chooses an action, $\boldsymbol{a}_i := (e_i, \boldsymbol{z}_i)$, which is combined with an idiosyncratic productivity shock, ε_i, to generate a random output, $y_i := y(\boldsymbol{a}_i) + \varepsilon_i$. Let

$$\begin{aligned} \pi_i := \pi(\boldsymbol{a}_i, \varepsilon_i) &:= y_i - \boldsymbol{p} \cdot \boldsymbol{z}_i \\ &= y(\boldsymbol{a}_i) + \varepsilon_i - \boldsymbol{p} \cdot \boldsymbol{z}_i \end{aligned} \tag{A.1}$$

be i's profit. Denote by Φ^{ε_i} and ϕ^{ε_i} the cumulative distribution function (CDF) and the probability density function (PDF) of ε_i. Letting $\widehat{\pi}_i$ be a realization of π_i, the CDF of π_i conditional on $\boldsymbol{a}_i$ is given by $\Phi^{\pi_i}(\widehat{\pi}_i \mid \boldsymbol{a}_i) := \Pr(\pi_i \leq \widehat{\pi}_i)$. This CDF can be calculated as follows:

$$\begin{aligned} \Phi^{\pi_i}(\widehat{\pi}_i \mid \boldsymbol{a}_i) &:= \Pr(\pi(\boldsymbol{a}_i; \varepsilon_i) \leq \widehat{\pi}_i) \\ &= \Pr(\varepsilon_i \leq \widehat{\pi}_i - y(\boldsymbol{a}_i) + \boldsymbol{p} \cdot \boldsymbol{z}_i) \\ &:= \Phi^{\varepsilon_i}(\widehat{\pi}_i - y(\boldsymbol{a}_i) + \boldsymbol{p} \cdot \boldsymbol{z}_i) \\ &= \int_{-\infty}^{\widehat{\pi}_i - y(\boldsymbol{a}_i) + \boldsymbol{p} \cdot \boldsymbol{z}_i} \phi^{\varepsilon_i}(\varepsilon_i)\, \mathrm{d}\varepsilon_i. \end{aligned} \tag{A.2}$$

This is a 'parametrized distribution representation' of profit. This representation highlights that different actions give rise to different distributions of profit. It turns out to be analytically convenient to work with both the parametrized distribution representation of profit and its primitive 'state-space representation,' given in Equation (A.1).[1] Given Equation (A.2), one can write

$$\begin{aligned}\phi^{\pi_i}\left(\widehat{\pi}_i \mid \boldsymbol{a}_i\right) &= \frac{\partial \Phi^{\pi_i}\left(\widehat{\pi}_i \mid \boldsymbol{a}_i\right)}{\partial \widehat{\pi}_i} \\ &= \frac{\mathrm{d}}{\mathrm{d}\widehat{\pi}_i} \int_{-\infty}^{\widehat{\pi}_i - y(\boldsymbol{a}_i) + \boldsymbol{p}\cdot\boldsymbol{z}_i} \phi^{\varepsilon_i}\left(\varepsilon_i\right)\, \mathrm{d}\varepsilon_i \\ &= \phi^{\varepsilon_i}\left(\widehat{\pi}_i - y\left(\boldsymbol{a}_i\right) + \boldsymbol{p}\cdot\boldsymbol{z}_i\right).\end{aligned}$$

Throughout, I assume that ϕ^{ε_i} is differentiable.

Let $\boldsymbol{\pi} := (\pi_i)_i$ be a profile of profits. The consumption received by i when $\boldsymbol{\pi}$ realizes is denoted $c_i(\boldsymbol{\pi})$. The feasibility constraint dictates that

$$\sum_{i\in N} c_i\left(\boldsymbol{\pi}\right) \leq \sum_{i\in N} \pi_i,$$

for each $\boldsymbol{\pi}$. Household i's utility is

$$u\left(c_i\left(\boldsymbol{\pi}\right)\right) - \kappa e_i,$$

where u is twice-continuously differentiable, strictly increasing, and strictly concave.

Let $\boldsymbol{a} := (\boldsymbol{a}_i)_i$ and $\Phi^{\boldsymbol{\pi}}\left(\boldsymbol{\pi} \mid \boldsymbol{a}\right) := \prod_i \Phi^{\pi_i}\left(\pi_i \mid \boldsymbol{a}_i\right)$. This is the cumulative distribution function of $\boldsymbol{\pi}$, because profits are independent between households once conditioning on actions taken. Finally, let $\boldsymbol{c}(\boldsymbol{\pi}) := (c_i(\boldsymbol{\pi}))_i$ be the sharing contract. An allocation is a pair $(\boldsymbol{c}(\boldsymbol{\pi}), \boldsymbol{a})$.

Full information. Assume that $\boldsymbol{a}$ is observed by the planner. In this case, there are no information frictions, so the planner can implement any action profile at no cost. Formally, the planner's problem is

$$\begin{aligned} &\max_{\boldsymbol{c}(\boldsymbol{\pi}),\boldsymbol{a}} \sum_{i\in N} \left\{\mathbb{E}_{\boldsymbol{\pi}}\left[u\left(c_i\left(\boldsymbol{\pi}\right)\right) \mid \boldsymbol{a}\right] - ke_i\right\}, \\ &\text{subject to } \sum_{i\in N} c_i\left(\boldsymbol{\pi}\right) = \sum_{i\in N} \pi_i,\ \forall \boldsymbol{\pi}. \end{aligned} \tag{A.3}$$

Notice that the feasibility constraint holds with equality. This is without loss of generality, as the constraint must bind at a solution to the problem. The following proposition characterizes the optimal sharing contract under full information. The proposition shows that

[1] See, e.g., Conlon (2009) for a discussion of the differences between state-space and parametrized distribution representations.

the first-order conditions for households' consumptions imply that the ratio of marginal utilities across any two households is constant across profit realizations. This is Borch's rule; i.e., the condition for perfect risk sharing (when the solution is interior).

Proposition A.1.1. *Under full information there is perfect risk sharing.*

The following claim characterizes the an optimal action profile.

Claim A.1.1. *Let $(\boldsymbol{c}^{\diamond}(\boldsymbol{\pi}), \boldsymbol{a}^{\diamond})$ be a solution to Problem (A.3). At $\boldsymbol{a}^{\diamond}$, (i) the social expected marginal benefit of effort is equal to its private marginal cost, and (ii) the marginal product of any intermediate q is equal to its price.*

The intuition behind this corollary is provided by Samuelson's rule for the optimal provision of public goods. The rule states that, at an optimum, the social marginal benefit of a public good equals the marginal cost of providing it. The key is to notice that, when households share profits, effort is a public good: an increase in effort on the part of j directly affects i's consumption. On the other hand, the condition for the optimal use of intermediate q is the standard optimality condition for a market-provided private good. This is because under profit sharing agricultural inputs remain private goods.

Private information. Next, assume that the action taken by i and the shock it receives are private to i. In this case, profits are publicly observable, noisy signals of actions taken. After observing the signals, the planner collects the profits realized and redistributes them to the households according to the sharing contract he designs. The planner takes into account that the households non-cooperatively choose an action profile given the sharing contract. Formally, the planner's problem is

$$\begin{aligned} &\max_{\boldsymbol{c}(\boldsymbol{\pi}),\boldsymbol{a}} \sum_{i\in N} \{\mathbb{E}_{\boldsymbol{\pi}}\left[u\left(c_i\left(\boldsymbol{\pi}\right)\right) \mid \boldsymbol{a}\right] - ke_i\} \\ &\text{subject to } \sum_i c_i\left(\boldsymbol{\pi}\right) = \sum_i \pi_i, \ \forall \boldsymbol{\pi}, \\ &\text{and } \boldsymbol{a}_i \in \arg\max_{\widetilde{\boldsymbol{a}}_i} \mathbb{E}_{\boldsymbol{\pi}}\left[u\left(c_i\left(\boldsymbol{\pi}\right)\right) \mid \widetilde{\boldsymbol{a}}_i, \boldsymbol{a}_{-i}\right] - k\widetilde{e}_i, \ \forall i. \end{aligned} \tag{A.4}$$

The IC constraints essentially define a pure-strategy Nash equilibrium: at $\boldsymbol{a}$, no household wants to deviate when it correctly anticipates the other households' actions. The set of IC constraints is a complicated object, as it comprises of a set of intertwined optimization problems. Moreover, there might exist more than one Nash equilibrium (or even none).[2]

[2]If expected utility is continuous and concave in own action, continuous in others' actions, and action sets are compact and convex, then there exists a Nash equilibrium. Notice that the action sets in this model are not compact, as they are unbounded from above. Moreover, households' expected utility is not necessarily concave in own action. Hence, the existence of a Nash equilibrium cannot be guaranteed.

Many papers in the principal-agent literature dealing with similar contracting problems have relied on the first-order approach (FOA), by which the agent's IC constraint is replaced by its first-order conditions. The optimal contract is then easily derived. The literature (Rogerson (1985) and Jewitt (1988)) has then focused on providing sufficient conditions under which the FOA is valid. My problem is different from the canonical principal-agent problem as there are n agents and each of them is choosing a multidimensional action. It is worthwhile to set the stage by characterizing the optimal sharing contract under the assumption that the FOA is valid. More formally, I begin by considering a relaxed version of Problem (A.4), where the IC constraints are replaced with the requirement that the action chosen by each household be a stationary point, given the actions chosen by the other households and the sharing contract. The key assumption is that a solution to the relaxed version of the problem is also a solution to Problem (A.4).[3]

Assumption A.1.1. *Let* $(\boldsymbol{c}^*(\boldsymbol{\pi}), \boldsymbol{a}^*)$ *be a solution to the following relaxed version of Problem (A.4):*

$$\begin{aligned}
&\max_{\boldsymbol{c}(\boldsymbol{\pi}),\boldsymbol{a}} \sum_{i\in N} \left\{\mathbb{E}_{\boldsymbol{\pi}}\left[u\left(c_i\left(\boldsymbol{\pi}\right)\right) \mid \boldsymbol{a}\right] - ke_i\right\} \\
&\textit{subject to } \sum_i c_i\left(\boldsymbol{\pi}\right) = \sum_i \pi_i, \ \forall \boldsymbol{\pi}, \\
&\int u\left(c_i\left(\boldsymbol{\pi}\right)\right) \phi_{e_i}^{\pi_i}\left(\pi_i \mid \boldsymbol{a}_i\right) \prod_{j\neq i} \phi^{\pi_j}\left(\pi_j \mid \boldsymbol{a}_j\right) \mathrm{d}\boldsymbol{\pi} = k, \forall i, \\
&\int u\left(c_i\left(\boldsymbol{\pi}\right)\right) \phi_{z_i^q}^{\pi_i}\left(\pi_i \mid \boldsymbol{a}_i\right) \prod_{j\neq i} \phi^{\pi_j}\left(\pi_j \mid \boldsymbol{a}_j\right) \mathrm{d}\boldsymbol{\pi} = 0, \ \forall i, \forall q.
\end{aligned} \tag{A.5}$$

$(\boldsymbol{c}^*(\boldsymbol{\pi}), \boldsymbol{a}^*)$ *is a solution to Problem (A.4).*

The following proposition characterizes the optimal sharing rule under private information when Assumption A.1.1 holds.

Proposition A.1.2. *Suppose that Assumption A.1.1 holds. Then, the optimal sharing rule is pinned down by*

$$\frac{u'\left(c_i^*\left(\boldsymbol{\pi}\right)\right)}{u'\left(c_j^*\left(\boldsymbol{\pi}\right)\right)} = \frac{1+\psi_j\Lambda_j\left(\pi_j \mid \boldsymbol{a}_j^*\right)}{1+\psi_i\Lambda_i\left(\pi_i \mid \boldsymbol{a}_i^*\right)}, \tag{A.6}$$

for each $i, j \in N$*, where* ψ_i *is the Lagrange multiplier associated to household* i*'s first-*

[3] Most likely, it is not so useful to give general conditions for the validity of the FOA, as these would probably not be conditions that generalize the specific assumptions of the model in Section 1.2. I would be better off showing that with quadratic utility (which implies mean-variance expected utility) and Lagrange shocks, the first-order approach is valid, as suggested in Wang (2013).

order condition for effort and

$$\Lambda_i\left(\pi_i \mid \boldsymbol{a}_i^*\right) := \frac{\phi_{e_i}^{\pi_i}\left(\pi_i \mid \boldsymbol{a}_i^*\right)}{\phi^{\pi_i}\left(\pi_i \mid \boldsymbol{a}_i^*\right)}.$$

Equation (A.6) is a modified Borch rule. In particular, if $\psi_i \neq 0$ or $\psi_j \neq 0$, then the households' marginal utilities are not equalized across profit realizations. The wedge between Equation (A.6) and the Borch rule is designed by the planner to take into account the effect of the sharing contract on the incentives to exert effort. On the other hand, the planner does *not* take into account the effect of the sharing contract on the use of agricultural inputs. This follows because the households are sharing profits—the revenues of production minus the monetary costs of agricultural inputs. Profit sharing implies that the effect of the sharing contract on the marginal benefit of an agricultural input is compensated by the effect of the contract on the marginal cost of that input; i.e., the sharing contract does not distort the incentives to use inputs purchased in the market. Finally, the wedge between Equation (A.6) and the Borch rule implies that risk sharing is generally imperfect under private information, as shown in the following corollary.

Corollary A.1.1. *Under private information risk sharing is imperfect.*

In the following, I consider the interaction between insurance, effort choices, and use of agricultural inputs. To do so, I focus on the implementation of a given action profile, and analyze how actions change when the sharing contract is perturbed. A more satisfying approach would be to jointly deriving an optimal action profile and the optimal sharing contract implementing it as a function of the parameters of the model—the utility cost of effort, the price of agricultural inputs, the variance of the production shock, and so on. Then, one could analyze how exogenous changes in these parameters jointly affect the sharing contract and the action profile, and thus keep track of the relationship between risk sharing and actions. I choose to follow the first approach because jointly deriving an optimal action in addition to the optimal contract that implements it is typically a very complex problem. Analyzing how actions change when the sharing contract is perturbed allows for significant tractability and is useful for practical applications.[4] To gain tractability, consider the case in which the optimal sharing contract is differentiable,[5]

[4]In fact, most of the papers in both the theoretical and applied literatures on the principal-agent problem focus predominantly on implementing a given action (Edmans & Gabaix (2011)).

[5]This approach is not entirely satisfactory, as $\boldsymbol{c}^*(\boldsymbol{\pi})$ is an endogenous object which was computed by means of point-wise maximization; hence, there is no a priori reason to expect $\boldsymbol{c}^*(\boldsymbol{\pi})$ to be differentiable. While not being rigorous, this approach is common practice (see e.g. Appendix B in Attanasio & Pavoni (2011)).

and define the slope of the contract at $\boldsymbol{\pi}$ for household i as

$$\frac{\partial c_i^*(\boldsymbol{\pi})}{\partial \pi_i}.$$

Intuitively, the slope of the contract measures the responsiveness of consumption to income. The smaller is the slope of the contract at $\boldsymbol{\pi}$, the higher is the insurance it provides at that profit realization. If the sharing contract is linear, than its slope is constant. Moreover, as shown in Subsection A.4, in a closed economy with no savings, the slope of a linear sharing contract coincides with the within estimator, δ^W. The following claim generalizes the intuition provided by the model of exogenous risk sharing in Section 1.2 by showing that, when the optimal sharing contract is differentiable, making the contract steeper (i.e., decreasing insurance) for household i induces the household to exert more effort.

Claim A.1.2. *Assume that the optimal sharing contract is differentiable. The higher is the slope of the contract at any $\boldsymbol{\pi}$, the higher is the effort provided.*

Let

$$p\left(c_i^*(\boldsymbol{\pi})\right) := \frac{k}{\int u'\left(c_i^*(\boldsymbol{\pi})\right)\frac{\partial c_i^*(\boldsymbol{\pi})}{\partial \pi_i}\mathrm{d}\Phi^{\varepsilon}(\boldsymbol{\varepsilon})},$$

be the 'effective' price of effort. This result is based on the fact that this price is decreasing in the slope of the contract. Next, I prove the main theorem, which extends the results of Theorem 1.2.1 to the case in which risk sharing is endogenous.

Theorem A.1.1. *Let $(\boldsymbol{c}^*(\boldsymbol{\pi}), \boldsymbol{a}^*)$ be a solution to the planner's problem under private information. Suppose that e_i and z_i^q are substitutes at $(p(c_i^*(\boldsymbol{\pi})), \boldsymbol{p})$; i.e.,*

$$\frac{\partial z_i^{q*}\left(p\left(c_i^*(\boldsymbol{\pi})\right), \boldsymbol{p}\right)}{\partial p\left(c_i^*(\boldsymbol{\pi})\right)} > 0.$$

Then,

$$\frac{\partial z_i^{q*}\left(p\left(c_i^*(\boldsymbol{\pi})\right), \boldsymbol{p}\right)}{\partial \frac{\partial c_i^*(\boldsymbol{\pi})}{\partial \pi_i}} < 0.$$

The signs of the latter two inequalities are reversed if e_i and z_i^q are complements at $(p(c_i^*(\boldsymbol{\pi})), \boldsymbol{p})$.

This theorem generalizes Theorem 1.2.1. In particular, it makes it clear that all the results obtained in Section 1.2 can be obtained in a model of endogenous risk sharing where the first-order approach is valid and the optimal sharing contract is differentiable. In the latter case, decreasing α would amount to making the sharing contract steeper.

A.1.1 Mean-Variance Expected Utility

Problem (A.4) is a complicated one, as the n incentive-compatibility constraints define a pure-strategy Nash equilibrium. In this subsection, I show that if households have mean-variance expected utility and the optimal sharing rule is linear, Problem (A.4) can be greatly simplified as each household's optimal action is independent of what the other households do.[6] Mean-variance expected utility can be justified in two cases. First, one can assume that the utility from consumption is quadratic; i.e.,

$$u\left(c_i\left(\boldsymbol{\pi}\right)\right) = c_i\left(\boldsymbol{\pi}\right) - \frac{\rho}{2} c_i\left(\boldsymbol{\pi}\right)^2 .$$

In this case, the expected utility of consumption takes a mean-variance specification, independently of the distribution of $c_i\left(\boldsymbol{\pi}\right)$. Alternatively, if the sharing contract is indeed linear and the production shocks are normally distributed,[7] then $c_i\left(\boldsymbol{\pi}\right)$ is also normally distributed, and the households have mean-variance expected utility when their utility from consumption is CARA; i.e.,

$$u\left(c_i\left(\boldsymbol{\pi}\right)\right) = -\exp\left\{-\rho c_i\left(\boldsymbol{\pi}\right)\right\} .$$

I begin by showing that, in the two cases justifying mean-variance expected utility, the optimal sharing contract under full information is equal sharing. Let $\boldsymbol{c}^{\diamond}$ be the optimal sharing rule under full information. The following claim holds:

Claim A.1.3. *If the households have quadratic utility from consumption, or if they have CARA utility from consumption and the planner can only use linear contracts, then* $c_i^{\diamond}\left(\boldsymbol{\pi}\right) = \overline{\pi}$, *for each* $i, j \in N$.

Mean-variance expected utility is particularly tractable because household i's choices are independent of the other households' choices when the sharing contract is linear. I proceed by demonstrating this result under the assumption that the sharing contract is linear. Let $\boldsymbol{c}^*$ be the optimal sharing rule under private information. Assume that $\boldsymbol{c}^*$ is linear; i.e.,

$$c_i^*\left(\boldsymbol{\pi}\right) = \sum_{j \in N} \alpha_{ij}^* \pi_j ,$$

for some $\left(\alpha_{ij}^*\right)_{ij} \in [0,1]^{2n}$ such that $\sum_i \alpha_{ij}^* = 1$, for each $j \in N$. The following claim proves that when the households have mean-variance expected utility, household i's choices are

[6] I.e., in this case, the n incentive-compatibility constraints define a dominant strategy equilibrium.

[7] The second argument *assumes* that the sharing contract is linear. In fact, if I were to posit that the sharing contract is chosen by the planner, then this argument would break, as Mirrlees famously shows that in a CARA-normal principal-agent model an optimal sharing contract does not exist (see Bolton & Dewatripont (2005)).

independent the other households' choices.

Claim A.1.4. *If $\mathbf{c}^*$ is linear and the households have mean-variance expected utility, household i's choices are independent of what the other households do.*

A.2 Effort and Inputs Affect the Variance of Output

Suppose that $y_i = y(a_i, \varepsilon_i)$, so that, in general,

$$\frac{\partial \mathbb{V}\text{ar}(y_i)}{\partial a_i} \neq \mathbf{0}.$$

In this case, given some α, household i's utility maximization problem is equivalent to

$$\max_{a_i} \left(1 - \frac{n-1}{n}\alpha\right) \mathbb{E}\left(y(a_i; \varepsilon_i) - pf_i\right) - \frac{\rho}{2} \mathbb{V}\text{ar}\left(y(a_i; \varepsilon_i)\right) - \kappa e_i.$$

Letting $\overline{y}(a_i) := \mathbb{E}(y(a_i; \varepsilon_i))$, the first-order conditions for e_i and f_i read

$$\overline{y}_e(a_i^*(\alpha)) = \frac{\kappa}{\left(1 - \frac{n-1}{n}\alpha\right)} + \frac{\rho}{2}\left(1 - \frac{n-1}{n}\alpha\right) \frac{\partial \mathbb{V}\text{ar}(y(a_i^*; \varepsilon_i))}{\partial e_i}$$

and

$$\overline{y}_f(a_i^*(\alpha)) = p + \frac{\rho}{2}\left(1 - \frac{n-1}{n}\alpha\right) \frac{\partial \mathbb{V}\text{ar}(y(a_i^*(\alpha); \varepsilon_i))}{\partial f_i}.$$

First of all, notice that if $\partial \mathbb{V}\text{ar}(y(a_i^*(\alpha); \varepsilon_i)) / \partial e_i < 0$ (i.e., supplying more effort reduces output volatility), then $\partial e_i^*(\alpha) / \partial \alpha < 0$, following the same steps used to prove Theorem 1.2.1. However, notice that now, contrary to what happens in the model outlined in Section 1.2, also the marginal cost of fertilizer is affected by risk sharing. More specifically, notice that if $\partial \mathbb{V}\text{ar}(y(a_i^*(\alpha); \varepsilon_i)) / \partial f_i < 0$ then risk sharing has direct negative effect on the use of fertilizer, while if $\partial \mathbb{V}\text{ar}(y(a_i^*(\alpha); \varepsilon_i)) / \partial f_i > 0$ then this direct effect is positive. This is intuitive: a direct effect of bettering insurance is that households optimally increase the use of inputs that make production riskier, while reducing the use of inputs that reduce output volatility. To this direct effect one has to sum the indirect effect coming from the complementarity between fertilizer and effort. In general, fertilizer is considered to be a risk-increasing technology. In general, one cannot say whether the direct effect will dominate the indirect one or vice-versa; however, to be sure, if the marginal impact of inputs on output volatility is sufficiently small, then Theorem 1.2.1 still holds.

A.3 Output Sharing

In this appendix, I show that the qualitative results of the model hold true when assuming that the households share outputs instead of profits. I will only analyze the private information regime, as one can easily show that, in the full information regime, the results of the model are the same under output sharing and profit sharing.

Suppose that household i's consumption is given by

$$c_i(\alpha) := (1-\alpha)\, y_i + \alpha \overline{y}.$$

Then, the welfare-maximizing action profile for a given α is characterized by

$$\begin{aligned} y_e\left(a_i^*(\alpha)\right) &= \frac{\kappa}{\left(1-\frac{n-1}{n}\alpha\right)} =: p_e, \\ \left(1-\frac{n-1}{n}\alpha\right) y_f\left(a_i^*(\alpha)\right) &= p, \end{aligned}$$

for each $i \in N$.

Let $p_e := \frac{\kappa}{\left(1-\frac{n-1}{n}\alpha\right)}$ and $p_f := \frac{p}{\left(1-\frac{n-1}{n}\alpha\right)}$. Effort and fertilizer are said to be complements at (p_e, p_f) if $\frac{\partial f_i^*(\alpha)}{\partial p_e} < 0$. If effort and fertilizer are complements, then

$$\frac{\partial f_i^*(\alpha)}{\partial \alpha} = \frac{\partial f_i^*(\alpha)}{\partial p_e}\frac{\partial p_e}{\partial \alpha} + \frac{\partial f_i^*(\alpha)}{\partial p_f}\frac{\partial p_f}{\partial \alpha} < 0.$$

A.4 Contrast Estimator

In this appendix, I describe the relationship between the risk-sharing contract specified in Equation (1.3) and the contrast estimator proposed by Suri (2011). Consider the following regression equation:

$$c_{it} - \overline{c}_{vt} = \delta^W\left(\pi_{it} - \overline{\pi}_{vt}\right) + \epsilon_{it}, \tag{A.7}$$

where c_{it} and π_{it} are household i's consumption and income in village and period t, and $\overline{c}_{vt}$ and $\overline{\pi}_{vt}$ are average consumption and income in village v and period t. Suri (2011) refers to δ^W as the within-estimator. Assume that the village is a closed economy with no saving technology. Then, the accounting identity $\overline{c}_{vt} = \overline{\pi}_{vt}$ trivially holds, and Equation (A.7) can be rewritten as follows:

$$c_{it} = \delta^W y_{it} + \left(1-\delta^W\right)\overline{y}_{vt} + \epsilon_{it}. \tag{A.8}$$

Equation (A.8) makes it clear that sharing rule α, as defined in Equation (1.3), theoretically coincides with δ^W when the village is a closed economy with no saving technology.

Next, consider the following regression equation:

$$\bar{c}_{vt} = \delta^B \bar{\pi}_{vt} + \epsilon_{vt}. \tag{A.9}$$

Suri (2011) refers to δ^B as the between-estimator, and defines the contrast estimator as follows:

$$\delta := 1 - \frac{\delta^W}{\delta^B}.$$

Note that if the village is a closed economy with no saving technology then $\delta^B = 1$. Thus, in this case, $\delta = 1 - \delta^W = 1 - \alpha$, and Equation (1.3) can be rewritten as follows:

$$c_{it} = \delta^W \pi_{it} + \delta \bar{\pi}_{vt}.$$

Table A12 reports the results of estimating Equation (A.7). The estimated coefficient is positive and significant. Table A13 reports the results of estimating Equation (A.9). A simple t-test reveals that $\widehat{\delta}^B$ is significantly lower than 1, suggesting the existence of inter-village risk sharing or saving technologies. Using point estimates $\widehat{\delta}^W$ and $\widehat{\delta}^B$ obtained in Tables A12 and A13, one can see that $\widehat{\delta} \approx 0.81$.

A.5 Complements and Substitutes

Theorems 1.2.1 and A.1.1 make use of the concepts of complementarity and substitutability. The theorems rely on the usual price-theoretic notion of complementarity and substitutability, as found for example in Mas-Colell et al. (1995). Here, I explore a different notion of complementarity and substitutability based on the concept of supermodularity, as in Acemoglu (2010), and show when the results of Theorems 1.2.1 and A.1.1 still apply under this other definition. For simplicity, I will restrict attention to the case in which the sharing contract is linear. The argument extends to the other case; just substitute α with the slope of the contract, as in Theorem A.1.1.

Acemoglu (2010) bases its definition of complementarity and substitutability on supermodularity. In particular, consider a two-input production function $y\left(e_i, z_i^q\right)$. Input q is strongly effort-complement if y is supermodular in (e_i, z_i^q), while it is strongly effort-saving if y is submodular in (e_i, z_i^q). These definitions are different but related to the price-theoretic definitions of complementarity and substitutability. Consider the following lemma:

Claim A.5.1. *If $y(e_i, z_i^q)$ is supermodular, then z_i^{q*} is decreasing in α. If $y(e_i, z_i^q)$ is submodular, then z_i^{q*} is increasing in α.*

If the production function can be written as

$$y(\boldsymbol{a}_i) = \sum_{q \in Q} y(e_i, z_i^q),$$

then Claim A.5.1 readily extents to the case in which there is more than one agricultural inputs.

A.6 Disentangling the Impact of Risk Sharing on Effort Supply and Fertilizer Use

Household i's problem can be written as

$$\max_{e_i, f_i} \ell_i^{1-\chi} \left[e_i^{\frac{\sigma-1}{\sigma}} + f_i^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\chi\sigma}{\sigma-1}} - pf_i - \frac{\kappa}{1 - \frac{n-1}{n}\alpha} e_i.$$

This is equivalent to the profit maximization problem of a competitive firm choosing effort and fertilizer while facing a real price of effort equal to $p_e = \kappa \left(1 - \frac{n-1}{n}\alpha\right)^{-1}$ and a real price of fertilizer equal to p. Since cost minimization is a necessary condition for profit maximization, consider the following cost minimization problem:

$$\begin{aligned} &\min_{e_i, f_i} pf_i + p_e e_i \\ &\text{subject to } \ell_i^{1-\chi} \left[e_i^{\frac{\sigma-1}{\sigma}} + f_i^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\chi\sigma}{\sigma-1}} \geq \widehat{y}_i \end{aligned}$$

Since land is fixed, the previous problem is equivalent to

$$\begin{aligned} &\min_{e_i, f_i} pf_i + p_e e_i \\ &\text{subject to } \left[e_i^{\frac{\sigma-1}{\sigma}} + f_i^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \geq y_i^{\dagger}, \end{aligned}$$

where $y_i^{\dagger} := \left(\frac{\widehat{y}_i}{\ell_i^{1-\chi}} \right)^{\frac{1}{\chi}}$. By the standard cost minimization problem with a CES technology, one obtains

$$e_i^*(\alpha) = \frac{y_i^{\dagger} {p_e}^{-\sigma}}{p^{1-\sigma} + {p_e}^{1-\sigma}}$$

and

$$f_i^*(\alpha) = \frac{y_i^{\dagger} p^{-\sigma}}{p^{1-\sigma} + {p_e}^{1-\sigma}}.$$

Taking logs, one gets

$$\log\left(e_i^*\left(\alpha\right)\right) = \log\left(y_i^\dagger\right) - \sigma\log\left(\kappa\right) + \sigma\log\left(1 - \frac{n-1}{n}\alpha\right) - \log\left(p^{1-\sigma} + \left(\frac{\kappa}{1-\frac{n-1}{n}\alpha}\right)^{1-\sigma}\right)$$

and

$$\log\left(f_i^*\left(\alpha\right)\right) = \log\left(y_i^\dagger\right) - \sigma\log\left(p\right) - \log\left(p^{1-\sigma} + \left(\frac{\kappa}{1-\frac{n-1}{n}\alpha}\right)^{1-\sigma}\right).$$

Using the structural estimates of σ and κ_i, and setting village size equal to the number of households sampled by ICRISAT, one can simulate the choices of fertilizer and effort for different levels of α. The only issue is that $y_i^\dagger$ is unobserved. To avoid this problem, I focus attention on the growth rates of effort and fertilizer when moving from α_0 to α_1. These growth rates can be calculated as follows:

$$\begin{aligned}\frac{e_{it}\left(\alpha_1\right) - e_{it}\left(\alpha_0\right)}{e_{it}\left(\alpha_0\right)} =& \log\left(e_{it}\left(\alpha_1\right)\right) - \log\left(e_{it}\left(\alpha_0\right)\right)\\ =&\widehat{\sigma}\log\left(\frac{\widetilde{n}_{vt}-1}{\widetilde{n}_{vt}}\alpha_1\right) - \log\left(\frac{\widetilde{n}_{vt}-1}{\widetilde{n}_{vt}}\alpha_1\right) +\\ &\log\left(p_{it}^{1-\widehat{\sigma}} + \left(\frac{\widehat{\kappa}_i}{1-\frac{\widetilde{n}_{vt}-1}{\widetilde{n}_{vt}}\alpha_0}\right)^{1-\widehat{\sigma}}\right) - \log\left(p_{it}^{1-\widehat{\sigma}} + \left(\frac{\widehat{\kappa}_i}{1-\frac{\widetilde{n}_{vt}-1}{\widetilde{n}_{vt}}\alpha_1}\right)^{1-\widehat{\sigma}}\right),\end{aligned}$$

and

$$\begin{aligned}\frac{f_{it}\left(\alpha_1\right) - f_{it}\left(\alpha_0\right)}{f_{it}\left(\alpha_0\right)} &= \log\left(f_{it}\left(\alpha_1\right)\right) - \log\left(f_{it}\left(\alpha_0\right)\right)\\ &= \log\left(p_{it}^{1-\widehat{\sigma}} + \left(\frac{\widehat{\kappa}_i}{1-\frac{\widetilde{n}_{vt}-1}{\widetilde{n}_{vt}}\alpha_0}\right)^{1-\widehat{\sigma}}\right) - \log\left(p_{it}^{1-\widehat{\sigma}} + \left(\frac{\widehat{\kappa}_i}{1-\frac{\widetilde{n}_{vt}-1}{\widetilde{n}_{vt}}\alpha_1}\right)^{1-\widehat{\sigma}}\right).\end{aligned}$$

Notice that these growth rates are independent of $y_i^\dagger$.

A.7 Computing the Optimal Sharing Rule

The aim is to solve Equation (1.6). To do so, I need an expression for $\partial e_i^*\left(\alpha\right)/\partial\alpha$. Let $P := \left({p_e}^{1-\sigma} + p^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$. The optimal use of fertilizer and effort are pinned down by

$$f_i^*\left(\alpha\right) = \left(\frac{p}{P}\right)^{-\sigma}\left(\frac{\chi}{P}\right)^{\frac{1}{1-\chi}}\ell_i$$

and

$$e_i^*\left(\alpha\right) = \left(\frac{p_e}{P}\right)^{-\sigma}\left(\frac{\chi}{P}\right)^{\frac{1}{1-\chi}}\ell_i$$

From this equation, I can compute $\partial e_i^* (\alpha) / \partial \alpha$. First, I make use of the chain rule to write

$$\frac{\partial e_i^* (\alpha)}{\partial \alpha} = \frac{\partial e_i^* (\alpha)}{\partial p_e} \frac{\partial p_e}{\partial \alpha}.$$

Then, notice that

$$\frac{\partial p_e}{\partial \alpha} = \frac{\kappa}{\left(1 - \frac{n-1}{n}\alpha\right)^2} \left(\frac{n-1}{n}\right)$$

and

$$\frac{\partial e_i^* (\alpha)}{\partial p_e} = \chi^{\frac{1}{1-\chi}} \ell_i \left[-\sigma {p_e}^{-\sigma-1} P^{\sigma - \frac{1}{1-\chi}} + {p_e}^{-2\sigma} \left(\sigma - \frac{1}{1-\chi}\right) P^{\sigma - \frac{1}{1-\chi} - 1} \left({p_e}^{1-\sigma} + p^{1-\sigma}\right)^{\frac{1}{1-\sigma} - 1} \right].$$

Finally,

$$\frac{\partial \mathbb{V}\mathrm{ar} (c_i (\alpha))}{\partial \alpha} = \left(-2 (1-\alpha) - \frac{2\alpha}{n} + \frac{2}{n}\right) \eta^2.$$

A.8 Computing the Optimal Fertilizer Subsidy

Assume that the government (i) aims to maximize utilitarian social welfare, (ii) can freely choose the price of fertilizer, and (iii) has no budget constraint. Then, the government's problem is equivalent to

$$\max_p \sum_{i \in N} \left[y (e_i^*, f_i^*) - p f_i^* - k e_i^* - \frac{\rho}{2} \mathbb{V}\mathrm{ar} (c_i (\alpha^*)) \right].$$

The first-order condition with respect to p reads

$$\sum_{i \in N} \left[y_e (e_i^*, f_i^*) \frac{\mathrm{d} e_i^*}{\mathrm{d} p} + y_f (e_i^*, f_i^*) \frac{\mathrm{d} f_i^*}{\mathrm{d} p} - f_i^* - p^r \frac{\mathrm{d} f_i^*}{\mathrm{d} p} - k \frac{\mathrm{d} e_i^*}{\mathrm{d} p} - \frac{\rho}{2} \frac{\partial \mathbb{V}\mathrm{ar} (c_i (\alpha^*))}{\partial \alpha^*} \frac{\partial \alpha^*}{\partial p} \right] = 0.$$

Notice that I apply the total derivative operator (with respect to p) to e_i^* and f_i^*. This is because e_i^* and f_i^* are functions of α^*, and α^* is itself a function of p. Collecting terms, one gets

$$\sum_{i \in N} \left[(y_e (e_i^*, f_i^*) - \kappa) \frac{\mathrm{d} e_i^*}{\mathrm{d} p} + (y_f (e_i^*, f_i^*) - p_r) \frac{\mathrm{d} f_i^*}{\mathrm{d} p} - \frac{\rho}{2} \frac{\partial \mathbb{V}\mathrm{ar} (c_i (\alpha^*))}{\partial \alpha^*} \frac{\partial \alpha^*}{\partial p} \right] = 0.$$

Using Claim 1.2.2, one obtains

$$\sum_{i \in N} \left[\kappa \left(\frac{1}{1 - \frac{n-1}{n}\alpha} - 1 \right) \frac{\mathrm{d} e_i^*}{\mathrm{d} p} - \frac{\rho}{2} \frac{\partial \mathbb{V}\mathrm{ar} (c_i (\alpha^*))}{\partial \alpha^*} \frac{\partial \alpha^*}{\partial p} \right] = 0.$$

Notice that

$$\frac{1}{1-\frac{n-1}{n}\alpha}-1>0$$

and $\mathrm{d}e_i^*/\mathrm{d}p<0$.

A.9 Proofs

Proof of Claim 1.2.1. Problem (1.4) is equivalent to

$$\max_{\boldsymbol{a}}\sum_{i\in N}\left((1-\alpha)\left(y\left(a_i\right)-pf_i\right)+\alpha\frac{\sum_{j\in N}y\left(a_j\right)-pf_j}{n}-\kappa e_i\right);$$

i.e.,

$$\max_{\boldsymbol{a}}\sum_{i\in N}\left((1-\alpha)\left(y\left(a_i\right)-pf_i\right)\right)+\alpha\sum_{j\in N}\left(y\left(a_j\right)-pz_j\right)-\sum_{i\in N}\kappa e_i.$$

If $\boldsymbol{a}^{\diamond}(\alpha)$ is an interior solution, then

$$(1-\alpha)\,y_e\left(a_k^{\diamond}(\alpha)\right)+\alpha y_e\left(a_k^{\diamond}(\alpha)\right)-\kappa=0,$$

for each $k\in N$; i.e., the marginal product of effort equals its marginal utility cost. The same argument holds for fertilizer. □

Proof of Claim 1.2.2. Problem (1.5) is equivalent to

$$\max_{a_i}\left(1-\frac{n-1}{n}\alpha\right)y\left(\left(a_i\right)-pf_i\right)-ke_i,\ \forall i\in N.$$

If $\boldsymbol{a}^*(\alpha)$ is an interior solution, then

$$\left(1-\frac{n-1}{n}\alpha\right)y_e\left(a_i^*(\alpha)\right)-k=0$$

and

$$\left(1-\frac{n-1}{n}\alpha\right)\left(y_f\left(a_i^*(\alpha)\right)-p\right)=0,$$

for each $i\in N$. □

Proof of Theorem 1.2.1. Notice that the household i's IC constraint is equivalent to the problem of a competitive firm facing a real price of fertilizer equal to p and a real price of effort equal to p_e. This is easily checked by considering the problem of such a firm and noticing that the profit-maximizing choices of effort and fertilizer coincide with the first-order conditions given in Claim 1.2.2. Notice that p_e is decreasing in α. By the law of supply, the demand for an input is decreasing in its price. Hence, $e_i^*(\alpha)$ is decreasing in α.

Moreover, α only affects p_e. Hence,

$$\begin{aligned}\frac{\partial f_i^*(\alpha)}{\partial \alpha} &= \frac{\partial f_i^*(\alpha)}{\partial p_e}\frac{\partial p_e}{\partial \alpha} \\ &= \frac{\partial f_i^*(\alpha)}{\partial p_e}\left(-\left(1-\frac{n-1}{n}\alpha\right)^{-1}\left(-\frac{n-1}{n}\right)\right).\end{aligned}$$

□

Proof of Claim 1.2.3. The problem of finding a welfare-maximizing sharing contract under full information is equivalent to

$$\max_{\alpha}\sum_{i\in N}\Bigg((1-\alpha)\left(y\left(a_i^{\diamond}(\alpha)\right)-pf_i^{\diamond}(\alpha)\right)+\alpha\frac{\sum_{j\in N}y\left(a_ij^{\diamond}(\alpha)\right)-pf_j^{\diamond}(\alpha)}{n}$$
$$-\frac{\rho}{2}\mathbb{V}\mathrm{ar}\left(c_i(\alpha)\right)-\kappa e_i^{\diamond}(\alpha)\Bigg),$$

where

$$\mathbb{V}\mathrm{ar}\left(c_i(\alpha)\right)=\left((1-\alpha)^2+\frac{\alpha^2}{n}+\frac{2\alpha(1-\alpha)}{n}\right)\eta^2.$$

Claim 1.2.1 implies that, under full information, $\boldsymbol{a}^{\diamond}(\alpha)$ is independent of α. Hence, the problem is equivalent to minimizing $\mathbb{V}\mathrm{ar}\left(c_i(\alpha)\right)$. It is easy to check that $\mathbb{V}\mathrm{ar}\left(c_i(\alpha)\right)$ is minimized when $\alpha=1$. □

Proof of Claim 1.2.4. The problem of finding a welfare-maximizing sharing contract under private information is equivalent to

$$\max_{\alpha}\sum_{i\in N}\left(\mathbb{E}\left(c_i(\alpha)\right)-\frac{\rho}{2}\mathbb{V}\mathrm{ar}\left(c_i(\alpha)\right)-\kappa e_i^*(\alpha)\right)$$

subject to

$$\begin{aligned}&\left(1-\frac{n-1}{n}\alpha\right)y_e\left(a_i^*(\alpha)\right)=\kappa,\\ &y_f\left(a_i^*(\alpha)\right)=p,\end{aligned}$$

for each $i\in N$. This problem can be written as

$$\max_{\alpha}\sum_{i\in N}\left(y\left(a_i^*(\alpha)\right)-pf_i^*(\alpha)+\mu-\kappa e_i^*(\alpha)\right)-\frac{n\rho}{2}\mathbb{V}\mathrm{ar}\left(c_i(\alpha)\right).$$

Derivate the planner's objective function with with respect to α to obtain

$$\sum_{i\in N}\left(y_e\left(a_i^*\left(\alpha\right)\right)\frac{\partial e_i^*\left(\alpha\right)}{\partial\alpha}+y_f\left(a_i^*\left(\alpha\right)\right)\frac{\partial f_i^*\left(\alpha\right)}{\partial\alpha}-p\frac{\partial f_i^*\left(\alpha\right)}{\partial\alpha}-\kappa\frac{\partial e_i^*\left(\alpha\right)}{\partial\alpha}\right)$$
$$-\frac{n\rho}{2}\frac{\partial\mathbb{V}\mathrm{ar}\left(c_i\left(\alpha\right)\right)}{\partial\alpha}.$$

Rearranging, I get

$$\sum_{i\in N}\left(\left(y_e\left(a_i^*\left(\alpha\right)\right)-\kappa\right)\frac{\partial e_i^*\left(\alpha\right)}{\partial\alpha}+\left(y_f\left(a_i^*\left(\alpha\right)\right)-p\right)\frac{\partial f_i^*\left(\alpha\right)}{\partial\alpha}\right)-\frac{n\rho}{2}\frac{\partial\mathbb{V}\mathrm{ar}\left(c_i\left(\alpha\right)\right)}{\partial\alpha}.$$

From the IC constraints given in Claim 1.2.2, the previous expression boils down to

$$\sum_{i\in N}\left(\kappa\left(\frac{1}{1-\frac{n-1}{n}\alpha}-1\right)\frac{\partial e_i^*\left(\alpha\right)}{\partial\alpha}\right)-\frac{n\rho}{2}\frac{\partial\mathbb{V}\mathrm{ar}\left(c_i\left(\alpha\right)\right)}{\partial\alpha}.$$

Notice that $\left(\frac{1}{1-\frac{n-1}{n}\alpha}-1\right)>0$, $\partial e_i^*\left(\alpha\right)/\partial\alpha<0$ by the law of supply (see the proof of Theorem 1.2.1), and $\partial\mathbb{V}\mathrm{ar}\left(c_i\left(\alpha\right)\right)/\partial\alpha<0$ (see the proof of Claim 1.2.3). □

Proof of Proposition A.1.1. Let $\left(\boldsymbol{c}^*\left(\boldsymbol{\pi}\right),\boldsymbol{a}^*\right)$ be a solution to the planner's problem, and $\varpi\left(\boldsymbol{\pi}\right)$ be the Lagrange multiplier of the feasibility constraint when profit profile $\boldsymbol{\pi}$ realizes. The first-order conditions for $c_i\left(\boldsymbol{\pi}\right)$ are

$$u'\left(c_i^*\left(\boldsymbol{\pi}\right)\right)\phi^{\boldsymbol{\pi}}\left(\boldsymbol{\pi}\mid\boldsymbol{a}^*\right)=\varpi\left(\boldsymbol{\pi}\right).$$

Combining this with the first-order conditions for $c_j\left(\boldsymbol{\pi}\right)$ yields

$$\frac{u'\left(c_i^*\left(\boldsymbol{\pi}\right)\right)}{u'\left(c_j^*\left(\boldsymbol{\pi}\right)\right)}=1,$$

for each i,j and each $\boldsymbol{\pi}$. That is, for each profit realization, consumption is adjusted so that the households' marginal utilities are equalized. □

Proof of Claim A.1.1. The first-order condition for e_i reads[8]

$$\sum_{j\in N}\int u\left(c_j^\diamond\left(\boldsymbol{\pi}\right)\right)\phi_{e_i}^{\pi_i}\left(\pi_i\mid\boldsymbol{a}_i^\diamond\right)\prod_{k\neq i}\phi^{\pi_k}\left(\pi_k\mid\boldsymbol{a}_k^\diamond\right)\mathrm{d}\boldsymbol{\pi}=k,$$

Notice that the right-hand side of this equation is i's private marginal cost of effort, while the right-hand size is the marginal increase in social welfare associated to a unitary

[8]In the following, $\frac{\partial\phi^{\pi_i}\left(\pi_i|\boldsymbol{a}_i\right)}{\partial x}:=\phi_x^{\pi_i}\left(\pi_i\mid\boldsymbol{a}_i\right)$.

increase in i's effort. On the other hand, the first-order condition for z_i^q is given by

$$\sum_{j\in N}\int u\left(c_i^{\diamond}\left(\boldsymbol{\pi}\right)\right)\phi_{z_i^q}^{\pi_i}\left(\pi_i \mid \boldsymbol{a}_i^{\diamond}\right)\prod_{k\neq i}\phi^{\pi_k}\left(\pi_k \mid \boldsymbol{a}_k^{\diamond}\right)\,\mathrm{d}\boldsymbol{\pi}=0, \tag{A.10}$$

Recall that $\phi^{\pi^i}\left(\pi_i \mid \boldsymbol{a}_i\right)=\phi^{\varepsilon^i}\left(\pi_i-y\left(\boldsymbol{a}_i\right)+\boldsymbol{p}\cdot\boldsymbol{z}_i\right)$. Hence,

$$\phi_{z_i^q}^{\pi_i^i}\left(\pi_i \mid \boldsymbol{a}_i\right)=\phi_{\varepsilon_i}^{\varepsilon^i}\left(\pi_i-y\left(\boldsymbol{a}_i\right)+\boldsymbol{p}\cdot\boldsymbol{z}_i\right)\left[-y_{z^q}\left(\boldsymbol{a}_i\right)+p^q\right].$$

Thus, Equation (A.10) can be rewritten as

$$\left[-y_{z^q}\left(\boldsymbol{a}_i^*\right)+p^q\right]\sum_{j\in N}\int u\left(c_j^{\diamond}\left(\boldsymbol{\pi}\right)\right)\phi_{\varepsilon_i}^{\varepsilon^i}\left(\pi_i-y\left(\boldsymbol{a}_i^{\diamond}\right)+\boldsymbol{p}\cdot\boldsymbol{z}_i^{\diamond}\right)\prod_{k\neq i}\phi^{\pi_k}\left(\pi_k \mid \boldsymbol{a}_k^{\diamond}\right)\,\mathrm{d}\boldsymbol{\pi}=0,$$

which is true if and only if $y_{z^q}\left(\boldsymbol{a}_i^{\diamond}\right)-p^q$. That is, the marginal product of intermediate q is equal to its price. □

Proof of Proposition A.1.2. Recall that $\phi^{\pi^i}\left(\pi_i \mid \boldsymbol{a}_i\right)=\phi^{\varepsilon^i}\left(\pi_i-y\left(\boldsymbol{a}_i\right)+\boldsymbol{p}\cdot\boldsymbol{z}_i\right)$. Hence,

$$\phi_{z_i^q}^{\pi_i^i}\left(\pi_i \mid \boldsymbol{a}_i\right)=\phi_{\varepsilon_i}^{\varepsilon^i}\left(\pi_i-y\left(\boldsymbol{a}_i\right)+\boldsymbol{p}\cdot\boldsymbol{z}_i\right)\left[-y_{z^q}\left(\boldsymbol{a}_i\right)+p^q\right].$$

As a consequence, the first-order condition for z_i^q can be rewritten as

$$\left[-y_{z^q}\left(\boldsymbol{a}_i\right)+p^q\right]\int u\left(c_i\left(\boldsymbol{\pi}\right)\right)\phi_{\varepsilon_i}^{\varepsilon^i}\left(\pi_i-y\left(\boldsymbol{a}_i\right)+\boldsymbol{p}\cdot\boldsymbol{z}_i\right)\prod_{j\neq i}\phi^{\pi_j}\left(\pi_j \mid \boldsymbol{a}_j\right)\,\mathrm{d}\boldsymbol{\pi}=0,$$

which is true as long as

$$y_{z^q}\left(\boldsymbol{a}_i\right)=p^q.$$

Hence, Problem (A.5) can be rewritten as

$$\begin{aligned}
&\max_{\boldsymbol{c}(\boldsymbol{\pi}),\boldsymbol{a}}\sum_{i\in N}\left\{\mathbb{E}_{\boldsymbol{\pi}}\left[u\left(c_i\left(\boldsymbol{\pi}\right)\right) \mid \boldsymbol{a}\right]-ke_i\right\}\\
&\text{subject to }\sum_i c_i\left(\boldsymbol{\pi}\right)=\sum_i \pi_i,\ \forall\boldsymbol{\pi},\\
&\int u\left(c_i\left(\boldsymbol{\pi}\right)\right)\phi_{e_i}^{\pi_i}\left(\pi_i \mid \boldsymbol{a}_i\right)\prod_{j\neq i}\phi^{\pi_j}\left(\pi_j \mid \boldsymbol{a}_j\right)\,\mathrm{d}\boldsymbol{\pi}=k,\forall i,\\
&y_{z^q}\left(\boldsymbol{a}_i\right)=p^q,\ \forall i,\forall q.
\end{aligned} \tag{A.11}$$

The Lagrangian associated to Problem (A.11) is

$$
\begin{aligned}
\mathcal{L}\left(\boldsymbol{c}\left(\boldsymbol{\pi}\right),\boldsymbol{a}\right) := \int \Bigg\{ & \sum_i \left[u\left(c_i\left(\boldsymbol{\pi}\right)\right) - ke_i\right] \\
& - \varpi\left(\boldsymbol{\pi}\right)\left[\sum_i c_i\left(\boldsymbol{\pi}\right) - \sum_i \pi_i\right]\frac{1}{\phi^{\boldsymbol{\pi}}\left(\boldsymbol{\pi} \mid \boldsymbol{a}\right)} \\
& - \sum_i \psi_i\left[u\left(c_i\left(\boldsymbol{\pi}\right)\right)\frac{\phi^{\pi_i}_{e_i}\left(\pi_i \mid \boldsymbol{a}_i\right)\prod_{j\neq i}\phi^{\pi_j}\left(\pi_j \mid \boldsymbol{a}_j\right)}{\phi^{\boldsymbol{\pi}}\left(\boldsymbol{\pi} \mid \boldsymbol{a}\right)} - v'\left(e_i\right)\right] \\
& - \sum_i\sum_q \xi_{iq}\left[y_{z^q}\left(\boldsymbol{a}_i\right) - p^q\right]\frac{1}{\phi^{\boldsymbol{\pi}}\left(\boldsymbol{\pi} \mid \boldsymbol{a}\right)}\Bigg\}\phi^{\boldsymbol{\pi}}\left(\boldsymbol{\pi} \mid \boldsymbol{a}\right)\,\mathrm{d}\boldsymbol{\pi}
\end{aligned}
$$

Then, the first-order condition for $c_i\left(\boldsymbol{\pi}\right)$ reads

$$
u'\left(c_i^*\left(\boldsymbol{\pi}\right)\right) - \frac{\varpi\left(\boldsymbol{\pi}\right)}{\phi^{\boldsymbol{\pi}}\left(\boldsymbol{\pi} \mid \boldsymbol{a}^*\right)} - \psi_i u'\left(c_i^*\left(\boldsymbol{\pi}\right)\right)\frac{\phi^{\pi_i}_{e_i}\left(\pi_i \mid \boldsymbol{a}_i^*\right)\prod_{j\neq i}\phi^{\pi_j}\left(\pi_j \mid \boldsymbol{a}_j^*\right)}{\phi^{\boldsymbol{\pi}}\left(\boldsymbol{\pi} \mid \boldsymbol{a}^*\right)} = 0. \tag{A.12}
$$

By independence, $\phi^{\boldsymbol{\pi}}\left(\boldsymbol{\pi} \mid \boldsymbol{a}\right) = \phi^{\pi_i}\left(\pi_i \mid \boldsymbol{a}_i\right)\prod_{j\neq i}\phi^{\pi_j}\left(\pi_j \mid \boldsymbol{a}_j\right)$. Hence, Equation (A.12) boils down to

$$
u'\left(c_i^*\left(\boldsymbol{\pi}\right)\right) - \frac{\varpi\left(\boldsymbol{\pi}\right)}{\phi^{\boldsymbol{\pi}}\left(\boldsymbol{\pi} \mid \boldsymbol{a}^*\right)} - \psi_i u'\left(c_i^*\left(\boldsymbol{\pi}\right)\right)\Lambda_i\left(\pi_i \mid \boldsymbol{a}_i^*\right) = 0,
$$

Combining this with the first-order condition for $c_j\left(\boldsymbol{\pi}\right)$ delivers Equation (A.6). □

Proof of Claim A.1.1. Perfect sharing requires Equation (A.6) to be constant across profit realizations. Suppose this is true; i.e.,

$$
\frac{1 + \psi_j\Lambda_j\left(\pi_j \mid \boldsymbol{a}_j^*\right)}{1 + \psi_i\Lambda_i\left(\pi_i \mid \boldsymbol{a}_i^*\right)} = r_{ij}, \tag{A.13}
$$

where r_{ij} is a constant. Rearrange Equation (A.13) to

$$
r_{ij}\psi_i\Lambda_i\left(\pi_i \mid \boldsymbol{a}_i^*\right) - \psi_j\Lambda_j\left(\pi_j \mid \boldsymbol{a}_j^*\right) = 1 - r_{ij} := \widehat{r}_{ij},
$$

where $\widehat{r}_{ij}$ is yet another constant. Multiply both sides of the previous equation by $\phi^{\pi_i}\left(\pi_i \mid \boldsymbol{a}_i^*\right)$ to obtain

$$
r_{ij}\psi_i\phi^{\pi_i}_{e_i}\left(\pi_i \mid \boldsymbol{a}_i^*\right) - \psi_j\frac{\phi^{\pi_j}_{e_j}\left(\pi_j \mid \boldsymbol{a}_j^*\right)}{\phi^{\pi_j}\left(\pi_j \mid \boldsymbol{a}_j^*\right)}\phi^{\pi_i}\left(\pi_i \mid \boldsymbol{a}_i^*\right) = \widehat{r}_{ij}\phi^{\pi_i}\left(\pi_i \mid \boldsymbol{a}_i^*\right).
$$

Integrate over π_i using the fact that $\int \phi^{\pi_i}(\pi_i \mid \boldsymbol{a}_i^*) = 1$ to get

$$r_{ij}\psi_i \int \phi_{e_i}^{\pi_i}(\pi_i \mid \boldsymbol{a}_i^*)\,\mathrm{d}\pi_i - \psi_j \frac{\phi_{e_j}^{\pi_j}\left(\pi_j \mid \boldsymbol{a}_j^*\right)}{\phi^{\pi_j}\left(\pi_j \mid \boldsymbol{a}_j^*\right)} = \widehat{r}_{ij}.$$

Next, multiply both sides of the previous equations by $\phi^{\pi_j}\left(\pi_j \mid \boldsymbol{a}_j^*\right)$, integrate over π_i, and use the fact that $\int \phi^{\pi_j}\left(\pi_j \mid \boldsymbol{a}_j^*\right) = 1$ to obtain

$$r_{ij}\psi_i \int \phi_{e_i}^{\pi_i}(\pi_i \mid \boldsymbol{a}_i^*)\,\mathrm{d}\pi_i - \psi_j \int \phi_{e_j}^{\pi_j}\left(\pi_j \mid \boldsymbol{a}_j^*\right)\,\mathrm{d}\pi_j = \widehat{r}_{ij}.$$

Notice that $\int \phi_{e_i}^{\pi_i}(\pi_i \mid \boldsymbol{a}_i^*)\,\mathrm{d}\pi_i = \int \phi_{e_j}^{\pi_j}\left(\pi_j \mid \boldsymbol{a}_j^*\right)\,\mathrm{d}\pi_j = 0$, since $\int \phi^{\pi_i}(\pi_i \mid \boldsymbol{a}_i^*) = \int \phi^{\pi_j}\left(\pi_j \mid \boldsymbol{a}_j^*\right) = 1$. Hence, it must be the case that

$$\widehat{r}_{ij} = 1 - r_{ij} = 0.$$

This is true if and only if $r_{ij} = 1$. Combining this last observation with Equation (A.13), one gets

$$\frac{1 + \psi_j \Lambda_j\left(\pi_j \mid \boldsymbol{a}_j^*\right)}{1 + \psi_i \Lambda_i(\pi_i \mid \boldsymbol{a}_i^*)} = 1;$$

i.e.,

$$\psi_j \Lambda_j\left(\pi_j \mid \boldsymbol{a}_j^*\right) = \psi_i \Lambda_i(\pi_i \mid \boldsymbol{a}_i^*). \tag{A.14}$$

Suppose $\psi_i, \psi_j \neq 0$ (otherwise perfect risk sharing would trivially obtain). Next, I show that Equation (A.14) cannot hold for each $\boldsymbol{\pi}$. To see this, pick $(\pi_j, \boldsymbol{\pi}_{-j}) = (\widehat{\pi}_j, \boldsymbol{\pi}_{-j})$. Equation (A.14) implies that

$$\Lambda_i(\pi_i \mid \boldsymbol{a}_i^*) = \frac{\psi_j}{\psi_i}\Lambda_j\left(\widehat{\pi}_j \mid \boldsymbol{a}_j^*\right).$$

Next, pick $(\pi_j, \boldsymbol{\pi}_{-j}) = \left(\widehat{\pi}'_j, \boldsymbol{\pi}_{-j}\right)$, with $\widehat{\pi}_j \neq \widehat{\pi}'_j$. Since Equation (A.14) holds for each $\boldsymbol{\pi}$, it must be the case that

$$\frac{\psi_j}{\psi_i}\Lambda_j\left(\widehat{\pi}_j \mid \boldsymbol{a}_j^*\right) = \Lambda_i(\pi_i \mid \boldsymbol{a}_i^*) = \frac{\psi_j}{\psi_i}\Lambda_j\left(\widehat{\pi}'_j \mid \boldsymbol{a}_j^*\right).$$

Given that the choices of $\widehat{\pi}_j$ and $\widehat{\pi}'_j$ were totally arbitrary, I conclude that $\Lambda_j\left(\pi_j \mid \boldsymbol{a}_j^*\right)$ must be a constant function of π_j. Hence, it must be the case that

$$\phi_{e_j}^{\pi_j}\left(\pi_j \mid \boldsymbol{a}_j^*\right) = w_j \phi^{\pi_j}\left(\pi_j \mid \boldsymbol{a}_j^*\right),$$

for some constant w_j. This is a first-order linear differential equation in e_i. The solution to this equation is given by

$$\phi^{\pi_j}\left(\pi_j \mid \boldsymbol{a}_j^*\right) = \frac{1}{\exp\left\{\int_0^E w_j \, \mathrm{d}e_i\right\}} \int_0^E \exp\left\{\int_0^E w_j \, \mathrm{d}e_i\right\} 0 \, \mathrm{d}e_i = 0.$$

This contradicts Equation (A.2). □

Proof of Claim A.1.2. Applying a change of variables from $\boldsymbol{\pi}$ to $\boldsymbol{\varepsilon}$ and assuming that the optimal sharing contract is differentiable, one can write the first-order condition for effort as

$$\int u'\left(c_i^*\left(\boldsymbol{\pi}\right)\right) \frac{\partial c_i^*\left(\boldsymbol{\pi}\right)}{\partial \pi_i} y_e\left(\boldsymbol{a}_i^*\right) \mathrm{d}\Phi^{\boldsymbol{\varepsilon}}\left(\boldsymbol{\varepsilon}\right) = k.$$

This can be rewritten as

$$y_e\left(\boldsymbol{a}_i^*\right) = p\left(c_i^*\left(\boldsymbol{\pi}\right)\right).$$

Notice that, when Assumption A.1.1 holds, household i's problem is equivalent to that of a competitive firm facing a real price of input q equal to p^q and a real price of effort equal to $p\left(c_i^*\left(\boldsymbol{\pi}\right)\right)$. By the law of supply, e_i^* is strictly decreasing in $p\left(c_i^*\left(\boldsymbol{\pi}\right)\right)$. Finally, notice that $p\left(c_i^*\left(\boldsymbol{\pi}\right)\right)$ is increasing in $\frac{\partial c_i^*(\boldsymbol{\pi})}{\partial \pi_i}$. □

Proof of Theorem A.1.1. By Claim A.1.2, e_i^* is increasing in $\frac{\partial c_i^*(\boldsymbol{\pi})}{\partial \pi_i}$. Notice that $\frac{\partial c_i^*(\boldsymbol{\pi})}{\partial \pi_i}$ affects $p\left(c_i^*\left(\boldsymbol{\pi}\right)\right)$, but not the prices of the other inputs. Hence,

$$\frac{\partial z_i^{q*}\left(p\left(c_i^*\left(\boldsymbol{\pi}\right)\right), \boldsymbol{p}\right)}{\partial \frac{\partial c_i^*(\boldsymbol{\pi})}{\partial \pi_i}} = \frac{\partial z_i^{q*}\left(p\left(c_i^*\left(\boldsymbol{\pi}\right)\right), \boldsymbol{p}\right)}{\partial p\left(c_i^*\left(\boldsymbol{\pi}\right)\right)} \frac{\partial p\left(c_i^*\left(\boldsymbol{\pi}\right)\right)}{\partial \frac{\partial c_i^*(\boldsymbol{\pi})}{\partial \pi_i}}.$$

□

Proof of Claim A.1.3. Proposition A.1.1 shows that, under full information, the optimal sharing rule is pinned down by the Borch rule:

$$\frac{u'\left(c_i^\diamond\left(\boldsymbol{\pi}\right)\right)}{u'\left(c_j^\diamond\left(\boldsymbol{\pi}\right)\right)} = 1,$$

If agents have quadratic utility from consumption, this boils down to

$$\frac{1 - \rho c_i^\diamond\left(\boldsymbol{\pi}\right)}{1 - \rho c_j^\diamond\left(\boldsymbol{\pi}\right)} = 1.$$

Hence, it must be the case that $c_i^\diamond\left(\boldsymbol{\pi}\right) = c_j^\diamond\left(\boldsymbol{\pi}\right)$, for each $j \neq i$. Using the feasibility constraint, this implies that $c_i^\diamond\left(\boldsymbol{\pi}\right) = \frac{\sum_j \pi_j}{n}$.

See Ambrus et al. (2017) for a proof that $c_i^\diamond\left(\boldsymbol{\pi}\right) = \overline{\pi}$ with CARA utility and linear contracts. □

Proof of Claim A.1.4. When $\boldsymbol{c}^*$ is linear and the households have mean-variance expected utility, household i's problem can be written as

$$\max_{\boldsymbol{a}_i} \sum_{j \in N} \alpha^*_{ij} \left(y\left(\boldsymbol{a}_j\right) - \boldsymbol{p} \cdot \boldsymbol{z}_j + \mu \right) - \frac{\rho}{2} \sum_{j \in N} {\alpha^*_{ij}}^2 \sigma^2 - \kappa e_i.$$

The objective function is continuously differentiable and jointly concave in e_i and z_i^q. Hence, the maximization problem is a concave program and the first-order conditions pin down an interior solution. The first-order conditions for e_i and z_i^q are given by

$$\alpha^*_{ii} y_e \left(\boldsymbol{a}^*_i\right) = \kappa$$

and

$$y_{z^q} \left(\boldsymbol{a}^*_i\right) = p^q,$$

respectively. Notice that these conditions are independent of $\boldsymbol{a}_j$, for $j \neq i$. □

Proof of Claim A.5.1. Recall that the household's problem is to that of a competitive firm facing a real price of input q equal to p^q and a real price of effort equal to p_e. Hence, household i's objective function can be written as

$$y\left(e_i, z_i^q\right) - p^q z_i^q - p_e e_i$$

Since $y\left(e_i, z_i^q\right)$ is increasing and supermodular, the household's objective function is supermodular in $(e_i, z_i^q, -p_e)$. Since the choice set is a sublattice, by Topkis' Monotonicity Theorem, $\left(e_i^*\left(p_e, p^q\right), z_i^{q*}\left(p_e, p^q\right)\right)$ is decreasing in p_e. Finally, notice that p_e is strictly increasing in α. This argument extends symmetrically for the case in which $y\left(e_i, z_i^q\right)$ is submodular. □

Appendix: Chapter 2

B.1 Further Details of the Empirical Analysis

Tables B.8–B.12 contains the regression tables of the directed networks on trust with the estimates of the coefficients associated to all the controls.

B.2 Instructions, Decision Sheet, and Surveys

B.2.1 First Stage

A

Código: ___________

¡Bienvenidos!

Algunos profesores de la Facultad de Economía estamos desarrollando una agenda de investigación que intenta entender como se forman las redes sociales, y qué factores dan lugar a dinámicas sociales más saludables, bajo las cuales podamos contribuir más, comunicarnos mejor y ser más creativos como individuos y como grupo. Para este fin los invitaremos periódicamente a participar en distintas actividades que nos permitirán medir algunos aspectos de su experiencia vital.

Esta es la primera actividad en la que los invitamos a participar.

En esta actividad les vamos a plantear un ejercicio de decisión y podrán ganar premios interesantes, por lo cual es importante que presten mucha atención. Es un ejercicio de decisión clásico de la economía, que ha sido replicado un gran número de veces en todo tipo de comunidades y con distintas variaciones. Mañana les contaremos un poco sobre la historia de este ejercicio.

La participación es totalmente voluntaria, y *la información que nos provean será mantenida de forma confidencial*. Más específicamente, esta información será utilizada solo para propósitos académicos, y para mejorar las estrategias de construcción de comunidades en la universidad. *Solo procesaremos y estudiaremos los datos una vez hayan sido anonimizados*.

<u>CONSENTIMIENTO INFORMADO</u>

Código de Uniandes ________________________________

Nombre y Apellidos ________________________________

e-mail ________________________________

Yo _________________________ con código Uniandes _________________________ declaro que fui informado de los ejercicios en que participaré en este taller, y acepto participar de manera voluntaria. Entiendo que mi información personal será mantenida de forma confidencial, y que será utilizada solo para propósitos académicos y para mejorar las estrategias de construcción de comunidades en la universidad.

Firma: _________________________ Fecha: _________________________

Yo, Tomás Rodríguez con cc.79982206 de Bogotá y profesor de la Universidad de Los Andes me comprometo a guardar la información de manera confidencial, a no revelarla a otros estudiantes y a utilizarla solo con fines académicos y de promoción de la construcción de comunidades en la universidad. Igualmente, me comprometo a entregar los premios que se asignen en cada caso para cada uno de los ejercicios de decisión. **Ni yo, ni los otros investigadores veremos sus identidades. Siempre trabajaremos exclusivamente con datos anónimos**. <u>En particular nosotros nunca veremos ni manipularemos esta hoja.</u>
<u>Sólo será usada por personal administrativo que estará encargado de hacerles llegar los premios y de fusionar información que ustedes nos provean en actividades posteriores.</u>

Firma: _________________________ Fecha: _________________________

B

Código: ___________

El Juego de las Transferencias

Este ejercicio se lleva a cabo en parejas de jugadores, en las que el jugador 1, llamado "el Remitente", y el jugador 2, llamado "el Receptor", tendrán roles distintos. Las parejas son anónimas; es decir, el Remitente no conoce, ni nunca conocerá, la identidad del Receptor con quien jugará. Del mismo modo, el Receptor no conoce, ni nunca conocerá, la identidad del Remitente con quien jugará.

En el juego, tanto el Remitente como el Receptor comienzan con $20.000 (en pesos colombianos) cada uno. El juego se desarrolla de la siguiente manera:

1. *El Remitente decide cuanto dinero enviarle al Receptor*. Tiene once posibilidades:

   $0, $2.000, $4.000, $6.000, $8.000, $10.00, $12.000, $14.000, $16.000, $18.000, $20.000.

2. *Al Receptor le llega tres veces la cantidad que el Remitente le haya enviado*. Es decir, si por ejemplo el Remitente decide enviarle $8.000 entonces le llegan

   $$3 \ x \ \$8.000 = \$24.000.$$

   (Si el Remitente le manda $0, le llegan $0.)

3. *El Receptor decide cuánto dinero*, del que haya recibido del Remitente, *le devuelve al Remitente*. Por ejemplo, si al Receptor le llegan $24.000, él o ella puede devolver al Remitente cualquier cantidad de dinero entre $0 y $24.000.

De esta manera, al final del juego, si x denota la cantidad de dinero que el Remitente decide enviarle al Receptor, y z la cantidad de dinero que el Receptor decide devolverle al Remitente,

el Remitente tendrá $\$20.000 - x + z$

y el Receptor tendrá $\$20.000 + 3x - z$.

Implementación del ejercicio: <u>Es fundamental que no hable , ni interactúe con ninguno de sus compañeros durante el ejercicio. Si tiene alguna duda por favor pregúntenos a nosotros</u>

(1) Le vamos a preguntar a cada uno de ustedes qué haría en caso de ser Remitente y qué haría en caso de ser Receptor.
(2) Un computador escogerá *al azar* quienes de ustedes jugarán como Remitentes y quienes como Receptores (mitad del grupo serán Remitentes y la otra mitad Receptores). También al azar los emaparejará (un remitente y un receptor en cada pareja) y con base en ese emparejamiento aleatorio calculará las ganancias finales de ustedes. **Noten que, por el diseño del ejercicio, ustedes nunca sabrán quien fue su pareja. Lo único que saben es que se trata de alguien más que está en este momento en este salón.**
(3) En una semana le haremos llegar un sobre a cada uno de ustedes con un bono equivalente a la suma que haya logrado en el juego. Este bono lo puede gastar en las tiendas Uniandes en lo que quieran.
<u>Recuerde: Para comenzar el juego le estamos realmente entregando $20.000 pesos, y cuánto dinero gane al final del ejercicio depende de las decisiones que tome (usted y su pareja).</u>

C

Código: ___________

Es fundamental que guarde silencio, durante el ejercicio. Si tiene alguna duda por favor pregúntenos a nosotros.

Un computador le asignará una pareja al azar y UNO solo de los dos roles posibles (en cada pareja habrá un Remitente y un Receptor).

Dado que de antemano no sabemos si usted será Remitente o Receptor, debemos preguntarle qué haría en cada caso.

NOTA IMPORTANTE: Dado lo anterior le garantizamos que usted NO jugará con usted mismo. Es decir, CON SEGURIDAD ABSOLUTA usted estará jugando con alguien más en este salón.

1. **Su decisión en caso de que le sea asignado el rol del jugador 1 (Remitente).**

   Por favor, marque con una *cruz* la cantidad de dinero que le enviaría al jugador 2 (Receptor). En la siguiente tabla, una unidad equivale a $2.000 pesos.

0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**

Por favor, marque UN solo número.

---

2. **Su decisión en caso de que le sea asignado el rol del jugador 2 (Receptor) en su pareja.**

   Para cada cantidad de dinero que posiblemente reciba del Remitente (esto depende de cuánto le envíe a usted el Remitente, y esto solo lo sabremos cuando el computador cree las parejas al azar), por favor, marque cuanto dinero le gustaría enviarle de vuelta. En el costado izquierdo de cada fila de la tabla que se presenta a continuación aparece una cantidad de dinero (una unidad equivale a $2.000) que el Remitente le puede haber enviado, y la cantidad que usted recibiría en cada caso (recuerde que recibirá tres veces la cantidad que el Remitente le envíe). En cada fila, por favor, marque con una *cruz* la cantidad de dinero que le enviaría de vuelta al Remitente, en ese caso. Note que en cada caso puede enviarle mínimo 0 y máximo la cantidad que haya recibido del Remitente. En la siguiente tabla, una unidad equivale a $2.000 pesos.

   Recuerde que además de lo que reciba del Remitente, usted, como Receptor, contará con los $20.000 que nosotros le damos. Pero no puede tocar sus $20.000; es decir, *solo puede enviarle dinero de vuelta al Remitente a partir de la cantidad que reciba gracias a lo que él o ella le transfiera*.

0	**→**	**0**	**0**																														
1	**→**	**3**	**0**	**1**	**2**	**3**																											
2	**→**	**6**	**0**	**1**	**2**	**3**	**4**	**5**	**6**																								
3	**→**	**9**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**																					
4	**→**	**12**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**																		
5	**→**	**15**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**															
6	**→**	**18**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**												
7	**→**	**21**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**									
8	**→**	**24**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**						
9	**→**	**27**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**			
10	**→**	**30**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**

Por favor verifique que haya marcado exactamente UN número en CADA fila.

D

Código ____________

1. Género (M, F)

2. Edad _____

3. ¿Cuántos hermanos(as) tiene? _____

4. Indique si está de acuerdo con las siguientes afirmaciones en una escala de 1 a 5. (Donde 5 representa "Totalmente de acuerdo" y 1, "Totalmente en Desacuerdo"

(a) No se puede confiar en los desconocidos. _____

(b) Cuando se trata desconocidos es mejor tener cuidado antes que confiar en ellos. _____

5.

(a) De sus 10 amigos más cercanos,
¿A cuántos les ha prestado dinero? _____

(b) De sus 10 amigos más cercanos,
¿Cuántos le han prestado dinero a usted? _____

(c) De sus 10 amigos más cercanos,
¿A cuántos les ha prestado sus pertenencias? _____
(i.e. Libros, CDs, ropa, bicicleta)

(d) De sus 10 amigos más cercanos,
¿Cuántos le han prestado a usted sus pertenencias? _____
(i.e. Libros, CDs, ropa, bicicleta)

(d) De sus 10 vecinos más cercanos,
¿A cuántos les confiaría las llaves de su casa? _____

(e) De sus 10 vecinos más cercanos,
¿Cuántos le confiarían a usted las llaves de sus casas? _____

6. ¿Cuántos amigos tiene fuera de la universidad? _____

7. ¿A cuánta gente de este grupo conocía antes de entrar a la universidad? _____

8. En general, usted diría que es :

A. Muy feliz

B. Bastante feliz

C. No muy feliz

D. Nada feliz

B.2.2 Second Stage

Introducción

¡Hola!
Gracias por participar en nuestra encuesta. En las siguientes ventanas, usted verá algunas preguntas. Por favor, respóndalas de forma intuitiva.
Tenga en cuenta que los datos recolectados durante esta encuesta serán usados únicamente con propósitos académicos. Usted permanecerá completamente anónimo ante tanto investigadores como el público.

Consentimiento informado

× SIGN HERE
clear

Por favor, escriba su código de estudiante.

Esta información no será revelada.

Browser Meta Info

This question will not be displayed to the recipient.
Browser: **Firefox**
Version: **66.0**
Operating System: **Ubuntu**
Screen Resolution: **1920x1080**
Flash Version: **32.0.0**
Java Support: **0**
User Agent: **Mozilla/5.0 (X11; Ubuntu; Linux x86_64; rv:66.0) Gecko/20100101 Firefox/66.0**

Browser Meta Info

This question will not be displayed to the recipient.
Browser: **Firefox**
Version: **66.0**
Operating System: **Ubuntu**
Screen Resolution: **1920x1080**
Flash Version: **32.0.0**
Java Support: **0**
User Agent: **Mozilla/5.0 (X11; Ubuntu; Linux x86_64; rv:66.0) Gecko/20100101 Firefox/66.0**

Conocimiento

En la siguiente ventana, usted encontrará una lista de personas de su cohorte. Seleccione las personas que usted saluda cuando se las encuentra

Por favor, tenga en cuenta que usted y sus amigos permanecerán anónimos: las respuestas estarán codificadas antes del análisis.

Por favor, seleccione si aplica

	Saludo a esta persona cuando me la encuentro
LINA MARIA ALVARADO	☐
JOSE GABRIEL PARDO	☐
MARIA PAULA ARDILA	☐
SOFIA VANEGAS	☐
MARIANA NOGUERA	☐
SANTIAGO GARCES	☐
FABIO EDUARDO RESTREPO	☐
NICOLAS GARCIA	☐
SAMUEL NARANJO	☐
LAURA LISETH GONZALEZ	☐
MARIA PAULA MEDINA	☐
JUAN DAVID PEREZ	☐
JUAN JOSE CHAVARRO	☐
ALEJANDRO PARRADO	☐

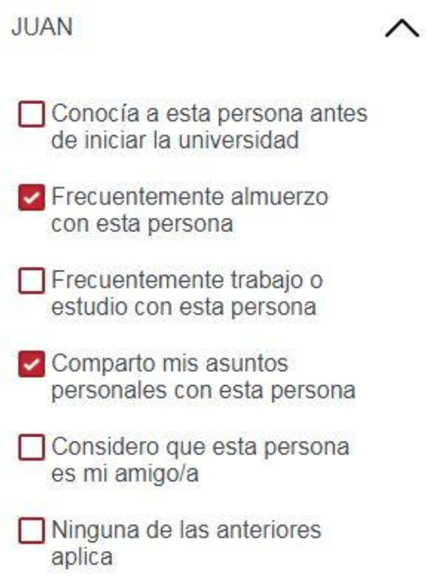
These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

Block 4

En la siguiente ventana, usted encontrará la lista de personas que usted indicó en la pregunta anterior. Debajo de cada uno de los nombres, encontrará una lista de afirmaciones.

Por favor, marque las afirmaciones que aplican a dicha persona. Marque TODAS las afirmaciones que apliquen.

Por ejemplo:



En la siguiente ventana, usted encontrará la lista de personas que usted indicó en la pregunta anterior. Debajo de cada uno de los nombres, encontrará una lista de afirmaciones.

Por favor, marque las afirmaciones que aplican a dicha persona.
Marque TODAS las afirmaciones que apliquen.

Por ejemplo:

	Conocía a esta persona antes de iniciar la universidad	Frecuentemente almuerzo con esta persona	Frecuentemente trabajo o estudio con esta persona	Comparto mis asuntos personales con esta persona	Considero que esta persona es mi amigo/a	Ninguna de las anteriores aplica
JUAN	☐	☑	☐	☑	☐	☐

En cada caso marque todas las afirmaciones que apliquen.

	Conocía a esta persona antes de iniciar la universidad	Frecuentemente almuerzo con esta persona	Frecuentemente trabajo o estudio con esta persona	Comparto mis asuntos personales con esta persona	Considero que esta persona es mi amigo/a	Ninguna de las anteriores aplica
» EDWIN DANIEL AGUIRRE	☐	☐	☐	☐	☐	☐
» NASIM AL ASHRAM	☐	☐	☐	☐	☐	☐
» LINA MARIA ALVARADO	☐	☐	☐	☐	☐	☐
» MAURICIO JOSE ARAGON	☐	☐	☐	☐	☐	☐
» JUAN SEBASTIAN ARCOS	☐	☐	☐	☐	☐	☐
» MARIA PAULA ARDILA	☐	☐	☐	☐	☐	☐
» JUAN SEBASTIAN AREVALO	☐	☐	☐	☐	☐	☐
» GERARDO ANDRES ARIAS	☐	☐	☐	☐	☐	☐

» MARIA ALESSANDRA

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

Preguntas Personales

¿Cuántos hermanos y hermanas tiene?

Por favor, escriba 0 si es hijo(a) único(a).

¿Cuántos amigos tiene que...

...frecuenten esta misma universidad, pero que conoció antes de entrar a la universidad?	0
...frecuenten esta misma universidad y que conoció en la universidad?	0
...no frecuenten esta misma universidad?	0
Total	0

¿Cuántas horas a la semana en promedio gasta socializando con amigos de la universidad?

¿Cuántas horas a la semana en promedio gasta socializando con amigos que no sean de la universidad?

¿Cuántas horas a la semana en promedio gasta en actividades físicas?

Por favor, liste sus pasatiempos (separados por comas).

Por favor, escriba la ciudad a la que usted considera su hogar.

Si usted considera varias ciudades como hogar, liste todas estas.

Por favor, seleccione las burbujas que mejor describan su personalidad.

Optimista	○	○	○	○	○	Realista
Extrovertida	○	○	○	○	○	Introvertida
Seguro	○	○	○	○	○	Cohibido
Animado	○	○	○	○	○	Tímido

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

Outlook

¿Cuántos años tiene?

¿Cuál es su color de ojos?

- Azul
- Café
- Verde
- Gris
- Negro
- Otro

¿Cuál es el color de su pelo?

- ○ Rubio
- ○ Castaño
- ○ Negro
- ○ Rojo
- ○ Otro

¿Cuál es su altura (en cm)?

¿Cuál es su peso (en kg)?

¿Usa gafas?

- Sí
- No

¿Tiene algún tatuaje?

- Sí
- No

¿Tiene piercings?

- Sí
- No

¿Usted fuma?

- No
- Ocasionalmente
- Todos los días

¿Asiste a fiestas?

- No
- De vez en cuando
- Frecuentemente

Por favor, seleccione las burbujas que mejor lo describan.

No soy sociable en lo absoluto	○	○	○	○	○	Soy muy sociable
Soy infeliz con mi vida social	○	○	○	○	○	Estoy satisfecho con mi vida social
Hacer amigos en la universidad es más dificil de lo que espeaba	○	○	○	○	○	Hacer amigos en la universidad es más facil de lo que esperaba
Me gustaría tener más amigos	○	○	○	○	○	Estoy satisfecho con el número de amigos que tengo

Seleccione la localidad donde vive

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

Muchas gracias por responder esta encuesta.
Por favor escribe una cuenta de correo a la cual podamos contactarte en caso de ganar el bono

Queremos escuchar sus sugerencias sobre esta encuesta.

B.3 Exponential Random Graph Models

The results of the logistic regressions presented in Section 2.3 might be biased because these regressions omit network characteristics as explanatory variables. For example, two individuals' decision to form a relationship between themselves might depend on whether they have friends in common, how many other relationships each of them is involved in, and other aspects of their social positions. If such interdependencies are relevant, estimation should not be conducted at the level of pairs of individuals but should encompass more of the network. ERGMs have become the workhorse models for estimating network formation thanks to their ability to incorporate general dependencies, but they are problematic because of their computational hurdle and since little is known about the consistency of estimators in such models. In any case, as a robustness check, we estimate several ERGMs. In table B.13 we present the results for one such ERGM when used to estimate the directed, the union, and the intersection of the friendship network. We include three measures of trust meant to capture generalized trust, particularized trust towards friends, and particularized trust towards neighbors. The first measure is obtained through a principal component analysis of the answers to questions A1, A2, and A3 (which we refer to as generalized trust); the second one is obtained through a principal component analysis of the answers to questions B1 and B3; and the third one is simply the answer to question C1. One can see that no trust measure has any significant effect on any of the networks. Results for other networks and ERGMs are in line with these findings.[1]

[1] Results are available upon request.

Appendix: Chapter 3

C.1 Proofs

Proof of Claim 3.2.1. If the principal does not run an appearance profile then the agents team up if and only if $(E^p u^1, E^p u^2) \geq 0$. By assumption, $E^p u^i \geq 0$. □

Proof of Claim 3.2.2. If the principal draws an appearance profile then the agents team up if and only if $(E^{p,\varepsilon} u^1(t^2), E^{p,\varepsilon} u^1(t^1)) \geq 0$. Notice that $E^{p,0} u^i(b) = u^i(\beta)$ and $E^{p,0} u^i(g) = u^i(\gamma)$. Thus, if $t = (g,g)$ then $E^{p,\varepsilon} u^i(t^j) > 0$, for each $i \in N$. On the other hand, for each $t \neq (g,g)$, there exists $i \in N$ such that $E^{p,\varepsilon} u^i(t^j) < 0$. □

Proof of Claim 3.2.3. First, notice that

$$E^{p,\varepsilon} u^i(g) = p_h^\gamma a + p_\ell^\beta (-c).$$

Since p_h^γ is strictly decreasing in ε and p_ℓ^β is strictly increasing in ε, $E^{p,\varepsilon} u^i(g)$ is strictly decreasing in ε. Then, notice that $E^p u^i = E^{p,\frac{1}{2}} u^i(g)$. □

Proof of Claim 3.2.4. First, notice that

$$E^{p,\varepsilon} u^i(b) = p_\ell^\gamma a + p_h^\beta (-c).$$

Since p_ℓ^γ is continuous and strictly increasing in ε and p_h^β is continuous and strictly decreasing in ε, $E^{p,\varepsilon} u^i(b)$ is continuous and strictly increasing in ε. Then, notice that $E^{p,0} u^i(b) = u^i(\beta) < 0$ and $E^{p,\frac{1}{2}} u^i(b) = E^p u^i \geq 0$. By the intermediate value theorem, there exists a point $\varepsilon^* \in \left(0, \frac{1}{2}\right]$ such that $E^{p,\varepsilon^*} u^i(b) = 0$. Since $E^{p,\varepsilon} u^i(b)$ is strictly increasing in ε, ε^* is unique. □

Proof of Claim 3.2.5. By Claim 3.2.4, $E^{p,\varepsilon} u^i(b) < 0$, for each $\varepsilon \in (0, \varepsilon^*)$. By Claim 3.2.3, $E^{p,\varepsilon} u^i(g) \geq 0$, for each $\varepsilon \in \left[0, \frac{1}{2}\right)$. Hence, $E^{p,\varepsilon} u^i(g) \geq 0$, for each $\varepsilon \in [0, \varepsilon^*)$. Since $E^{p,0} u^i(b) = u^i(\beta) < 0$, if $\varepsilon \in [0, \varepsilon^*)$, then the agents team up if and only if $t = (g,g)$.

By Claim 3.2.4, $E^{p,\varepsilon}u^i(b) > 0$, for each $\varepsilon \in \left[\varepsilon^*, \frac{1}{2}\right)$. By Claim 3.2.3, $E^{p,\varepsilon}u^i(g) \geq 0$, for each $\varepsilon \in (0, \varepsilon^*)$. Since $E^{p,\frac{1}{2}}u^i(b) = E^p u^i \geq 0$, $E^{p,\varepsilon}u^i(t^j) \geq 0$, for each $\varepsilon \in \left[\varepsilon^*, \frac{1}{2}\right]$ and each $t^j \in T^j$. That is, if $\varepsilon \in \left[\varepsilon^*, \frac{1}{2}\right]$ then the agents team up, for each $t \in T$. □

Proof of Claim 3.2.6. By Claim 3.2.5, if $\varepsilon \in \left[\varepsilon^*, \frac{1}{2}\right]$, then $\nVdash^{p,\varepsilon}_{team}(t) = \nVdash^{p,\frac{1}{2}}_{team}(t) = \nVdash^{p}_{team}(t)$, for each $t \in T$. □

Proof of Lemma 3.2.1. Assume that $\varepsilon \in [0, \varepsilon^*)$ and the principal draws an appearance profile. By Claim 3.2.5, the agents team up if and only if $t = (g, g)$. When $\theta = (\gamma, \gamma)$ (which happens with probability p^2), the probability that $t = (g, g)$ is $(1-\varepsilon)^2$. In this case, social welfare is $2a$. When $\theta \in \{(\gamma, \beta), (\beta, \gamma)\}$ (which happens with probability $2p(1-p)$), the probability that $t = (g, g)$ is $\varepsilon(1-\varepsilon)$. In this case, social welfare is $a - c$. When $\theta = (\beta, \beta)$ (which happens with probability $(1-p)^2$), the probability that $t = (g, g)$ is ε^2. In this case, social welfare is $-2c$. Finally, notice that expected social welfare when the principal does not draw an appearance profile is $p^2(2a) + 2p(1-p)(a-c) + (1-p)^2(-2c)$. □

Proof of Corollary 3.2.1. Given Lemma 3.2.1, one can show that

$$\frac{\partial^2 G^{p,\varepsilon}}{\partial \varepsilon^2} = -4(2p-1)(p(c-a)-c).$$

Notice that, since $a > c$, $p(c-a) - c < 0$. □

Proof of Corollary 3.2.2. One can show that $\widetilde{\varepsilon} = \frac{p^2(c-3a)+p(a-c)}{2(2p-1)(p(c-a)-c)}$. Notice that $p(c-a) - c < 0$, $p^2(c-3a) + p(a-c) \geq 0$ if and only if $p \leq \frac{c-a}{c-3a}$, and $\frac{c-a}{c-3a} < \frac{1}{2}$, so that $p \leq \frac{c-a}{c-3a}$ implies that $p < \frac{1}{2}$. Thus, if $p < \frac{c-a}{c-3a}$ then $\widetilde{\varepsilon} > 0$; if $p \geq \frac{c-a}{c-3a}$ and $p < \frac{1}{2}$, then $\widetilde{\varepsilon} \leq 0$; and if $p > \frac{1}{2}$ then $\widetilde{\varepsilon} > 0$. Finally, recall that $\varepsilon^* = \frac{(1-p)c}{pa+(1-p)c}$, and notice that if $p > \frac{1}{2}$ then $\widetilde{\varepsilon} > \varepsilon^*$. □

Proof of Corollary 3.2.3. Given Lemma 3.2.1, one can show that

$$\begin{aligned} G^{p,0} &= p^2 2a - \left(p^2 2a + 2p(1-p)(a-c) + (1-p)^2(-2c)\right) \\ &= p^2 a - p(a+c) + c \\ &= \left(p - \frac{c}{a}\right)(p-1). \end{aligned}$$

□

Proof of Corollary 3.2.4. Recall that $\varepsilon^* = \frac{(1-p)c}{pa+(1-p)c}$. Given Lemma 3.2.1, one can show that

$$\lim_{\eta \to 0} G^{p,\varepsilon^*-\eta} = \frac{-2(p-1)p(a+c)(p(c-a)-c)(p(a+c)-c)}{(pa+(1-p)c)^2}.$$

Notice that $p(c-a)-c<0$. $\square$

Proof of Theorem 3.2.1. By Claim 3.2.6, $\varepsilon \in \left[\varepsilon^*, \frac{1}{2}\right]$, $G^{p,\varepsilon}=0$. By Lemma 3.2.1, $G^{p,\varepsilon}$ is a degree 2 polynomial on $[0,\varepsilon^*)$. From Corollary 3.2.1, if $p>\frac{1}{2}$ then $G^{p,\varepsilon}$ is strictly convex on $[0,\varepsilon^*)$, if $p=\frac{1}{2}$ then $G^{p,\varepsilon}$ is linear on $[0,\varepsilon^*)$, and if $p<\frac{1}{2}$ then $G^{p,\varepsilon}$ is strictly concave on $[0,\varepsilon^*)$. This shows that $G^{p,\varepsilon}$ is smooth on $[0,\varepsilon^*)$.
Assume $p \geq \frac{c-a}{c-3a}$. Given Corollary 3.2.2, this shows that $G^{p,\varepsilon}$ is strictly decreasing on $[0,\varepsilon^*)$, whenever $p \neq \frac{1}{2}$. Notice that $\frac{c}{a} > \frac{c}{a+c}$. The rest follows from Corollaries 3.2.3 and 3.2.4.
Assume $p < \frac{c-a}{c-3a}$. Since $p<\frac{1}{2}$, $G^{p,\varepsilon}$ is strictly concave on $[0,\varepsilon^*)$. The rest directly follows from Corollary 3.2.2. $\square$

Proof of Claim 3.2.7. By Claims 3.2.3 and 3.2.4, if $\varepsilon \in \left[\varepsilon^*, \frac{1}{2}\right]$, then $E^{p,\varepsilon}u^i\left(t^i\right) \geq 0$, for each $t^i \in t^i$. Thus, the principal commands the agents to form a team independently of the appearance profile. Hence,

$$\begin{aligned} E^{p\,\varepsilon}DB &= \sum_{t\in T} \Pr(t) \sum_{i=1}^{2} E^{p,\varepsilon}u^i\left(t^j\right) \\ &= \sum_{i=1}^{2} E^p u^i \\ &= E^{p,\frac{1}{2}}SW. \end{aligned}$$

By Claim 3.2.3 and 3.2.4, if $\varepsilon \in [0,\varepsilon^*)$, $E^{p,\varepsilon}u^i(g)>0$ and $E^{p,\varepsilon}u^i(b;j)<0$. If $t=(g,g)$ then $\sum_{i=1}^{2} E^{p,\varepsilon}u^i\left(t^j\right)>0$. If $t\in\{(b,g),(g,b)\}$ then

$$\begin{aligned} \sum_{i=1}^{2} E^{p,\varepsilon}u^i\left(t^j;j\right) = \quad & p^2\left((1-\varepsilon)\varepsilon\right)2a + \\ & p(1-p)\left((1-\varepsilon)^2+\varepsilon^2\right)(a-c) + \\ & (1-p)^2\left((1-\varepsilon)\varepsilon\right)(-2c). \end{aligned}$$

If $\varepsilon=0$ then $\sum_{i=1}^{2} E^{p,\varepsilon}u^i\left(t^j;j\right)=a-c\geq 0$. Next, notice that

$$\begin{aligned} \frac{\partial \sum_{i=1}^{2} E^{p,\varepsilon}u^i\left(t^j;j\right)}{\partial \varepsilon} = \quad & p^2 2a - p^2 2a2\varepsilon - \\ & p(1-p)2(a-c) + 4p(1-p)(a-c)\varepsilon + \\ & (1-p)^2(-2c) - (1-p)^2(-2c)2\varepsilon. \end{aligned}$$

Rearranging, one gets

$$\frac{\partial \sum_{i=1}^{2} E^{p,\varepsilon} u^i \left(t^j; j\right)}{\partial \varepsilon} = \left(p^2 2a - 2p\left(1-p\right)\left(a-c\right) + \left(1-p\right)^2\left(-2c\right)\right) \cdot \left(1-2\varepsilon\right).$$

Notice that $1-2\varepsilon \geq 0$ and that

$$\left(p^2 2a - 2p\left(1-p\right)\left(a-c\right) + \left(1-p\right)^2\left(-2c\right)\right) \begin{cases} >0 & \text{if } p > \frac{1}{2} \\ =0 & \text{if } p = \frac{1}{2} \\ <0 & \text{if } p < \frac{1}{2} \end{cases}.$$

Thus, $E^{p,\varepsilon}$ is strictly decreasing in $[0, \varepsilon^*)$. □

Proof of Claim 3.2.7. If the principal does not draw an appearance profile then the agents team up if and only if $(E^p u^1, E^p u^2) \geq 0$. By assumption, $E^p u^i < 0$. □

Proof of Claim 3.2.8. Recall that

$$E^{p,\varepsilon} u^i \left(b\right) = p_\ell^\gamma a + p_h^\beta \left(-c\right).$$

Since p_ℓ^γ is strictly increasing in ε and p_h^β is strictly decreasing in ε, $E^{p,\varepsilon} u^i\left(b\right)$ is strictly increasing in ε. Notice that $E^p u^i\left(j\right) = E^{p,\frac{1}{2}} u^i \left(b\right)$. □

Proof of Claim 3.2.9. Recall that

$$E^{p,\varepsilon} u^i \left(g\right) = p_h^\gamma a + p_\ell^\beta \left(-c\right).$$

Since p_h^γ is continuous and strictly decreasing in ε and p_ℓ^β is continuous and strictly increasing in ε, $E^{p,\varepsilon} u^i\left(g\right)$ is continuous and strictly decreasing in ε. Notice that $E^{p,0} u^i \left(g\right) = u^i\left(\gamma\right) > 0$ and $E^{p,\frac{1}{2}} u^i\left(g\right) = E^p u^i < 0$. By the intermediate value theorem, there exists a point $\varepsilon^* \in \left(0, \frac{1}{2}\right)$ such that $E^{p,\varepsilon^*} u^i\left(g\right) = 0$. Since $E^{p,\varepsilon} u^i\left(g\right)$ is a strictly decreasing function of ε, ε^* is unique. □

Proof of Claim 3.2.10. By Claim 3.2.9, $E^{p,\varepsilon} u^i\left(g\right) \geq 0$, for each $\varepsilon \in \left(0, \varepsilon^*\right]$. By Claim 3.2.8, $E^{p,\varepsilon} u^i\left(b\right) < 0$, for each $\varepsilon \in \left[0, \frac{1}{2}\right]$. Since $E^{p,0} u^i\left(b\right) = u^i\left(b\right) < 0$, there exists $i \in N$ such that $E^{p,\varepsilon} u^i\left(b\right) < 0$, for each $\varepsilon \in [0, \varepsilon^*]$ and each $t \neq (g,g)$. That is, if $\varepsilon \in [0, \varepsilon^*)$, then the agents team up if and only if $t = (g,g)$.
By Claim 3.2.9, $E^{p,\varepsilon} u^i\left(g\right) < 0$, for each $\varepsilon \in \left(\varepsilon^*, \frac{1}{2}\right)$. By Claim 3.2.8, $E^{p,\varepsilon} u^i\left(b\right) < 0$, for each $\varepsilon \in \left[\varepsilon^*, \frac{1}{2}\right]$. Since $E^{p,\frac{1}{2}} u^i\left(g\right) = E^p u^i < 0$, $E^{p,\varepsilon} u^i\left(t^j\right) < 0$, for each $\varepsilon \in \left(\varepsilon^*, \frac{1}{2}\right]$ and $t^j \in T^j$. □

Proof of Claim 3.2.11. By Claim 3.2.10, if $\varepsilon \in \left[\varepsilon^*, \frac{1}{2}\right]$, then $\not\Vdash^{p,\varepsilon}_{team}(t) = \not\Vdash^{p,\frac{1}{2}}_{team}(t) = \not\Vdash^{p}_{team}(t)$, for each $t \in T$. □

Proof of Lemma 3.2.2. Assume that $\varepsilon \in [0, \varepsilon^*)$ and the principal draws an appearance profile. By by Claim 3.2.10, the agents team up if and only if $t = (g, g)$. When $\theta = (\gamma, \gamma)$ (which happens with probability p^2), the probability that $t = (g, g)$ is $(1-\varepsilon)^2$. In this case, social welfare is $2a$. When $\theta \in \{(\gamma, \beta), (\beta, \gamma)\}$ (which happens with probability $2p(1-p)$), the probability that $t = (g, g)$ is $\varepsilon(1-\varepsilon)$. In this case, social welfare is $a - c$. When $\theta = (\beta, \beta)$ (which happens with probability $(1-p)^2$), the probability that $t = (g, g)$ is ε^2. In this case, social welfare is $-2c$. By Claim 3.2.7, if the principal does not draw an appearance profile then the agents do not team up. Thus, $E^{p,\frac{1}{2}}SW = 0$, so that $G^{p,\varepsilon} = E^{p,\varepsilon}SW$. □

Proof of Lemma 3.2.3. Notice that for each $\varepsilon \in [0, \varepsilon^*]$,

$$\begin{aligned}
G^{p,\varepsilon} &= \sum_{\theta \in \Theta} \Pr(\theta) \sum_{t \in T} \Pr(t \mid \theta) \not\Vdash^{p,\varepsilon}_{team}(t) \sum_{i=1}^{2} u^i\left(\theta^j\right) \\
&= \sum_{\theta \in \Theta} \Pr(\theta) \sum_{t \in T} \Pr(t = (g,g) \mid \theta) \sum_{i=1}^{2} u^i(\gamma, \gamma) \\
&= \Pr(t = (g,g)) \left(\sum_{i=1}^{2} E^{p,\varepsilon} u^i(g) \right),
\end{aligned}$$

which is clearly positive. □

Proof of Corollary 3.2.5. Notice that if $p < \frac{c}{a+c}$ then $p < \frac{1}{2}$. The rest of the corollary follows from Corollary 3.2.1. □

Proof of Corollary 3.2.6. Notice that if $p < \frac{c}{a+c}$ then $p < \frac{1}{2}$. The rest of the result follows directly from Corollary 3.2.2. □

Proof of Corollary 3.2.7. Given Lemma 3.2.2, one can show that $G^{p,0} = p^2 2a$. □

Proof of Corollary 3.2.8. Recall that $\varepsilon^* = \frac{pa}{p(a-c)+c}$. Given Lemma 3.2.2, one can show that

$$\lim_{\eta \to 0} G^{p,\varepsilon^* - \eta} = \frac{p^2(1-p)^2 c^2 2a + 2p^2(1-p)^2 ac(a-c) + p^2(1-p)^2 a^2(-2c)}{(p(a-c)+c)^2}.$$

Thus, $\lim_{\eta \to 0} G^{p,\varepsilon^* - \eta} = 0$ if and only if

$$p^2(1-p)^2 c^2 2a + 2p^2(1-p)^2 ac(a-c) + p^2(1-p)^2 a^2(-2c) = 0,$$

which is always true. □

Proof of Theorem 3.2.2. By Claim 3.2.11, $\varepsilon \in \left[\varepsilon^*, \frac{1}{2}\right]$, $G^{p,\varepsilon} = 0$. By Lemma 3.2.2, $G^{p,\varepsilon}$ is a degree 2 polynomial on $[0, \varepsilon^*)$. From Corollary 3.2.5, $G^{p,\varepsilon}$ is strictly concave on $[0, \varepsilon^*)$. This shows that $G^{p,\varepsilon}$ is smooth on $[0, \varepsilon^*)$.

Assume that $p \geq \frac{c-a}{c-3a}$. Given Corollary 3.2.6, this shows that $G^{p,\varepsilon}$ is strictly decreasing on $[0, \varepsilon^*)$. The rest follows from Corollaries 3.2.7 and 3.2.8.

Assume that $p < \frac{c-a}{c-3a}$. Given Lemma 3.2.3, $G^{p,\varepsilon}$ is strictly positive on $\varepsilon \in [0, \varepsilon^*)$. The rest directly follows from Corollary 3.2.6. □