UNIVERSITÀ CATTOLICA
DEL SACRO CUORE
MILANO

Doctoral School in Economics

S.S.D: SECS-P/01

Essays on Development Economics

Valentina Rotondi
Department of Economics
PhD Matriculation Number: 4108781

a.a. 2014–15

Essays on Development Economics

Valentina Rotondi
Department of Economics
PhD Matriculation Number: 4108781

Supervisors
Luca Stanca
DEMS, University of Milan-Bicocca
Giuseppe Attanasi
BETA, University of Strasbourg

Doctoral School in Economics
S.S.D: SECS-P/01
PhD Director: **Gianluca Femminis**
a.a. 2014–15

Abstract

This dissertation makes use of several methodologies to explore topics ascribed to the field of development economics. Chapter 1 reviews the literature on social capital by presenting a decomposition of trust and networks – the cognitive and the structural component of social capital, respectively–, in several sub-dimensions. One of this dimension is used in chapter 2 where we investigate, both theoretically and empirically, the role played by the cultural norm of particularism, as opposed to universalism, for collusive bribery. Consistent with the theory, particularism is found to have a positive causal effect on the probability of offering a bribe. Chapter 3 assesses the impact of a small-scale agricultural extension project implemented in rural Ethiopia aimed at introducing the cultivation of horticultural gardens. Empirically, a mixed impact evaluation design is used combining across-villages comparisons, through difference-in-differences estimations, with a within village randomized control trial. The findings indicate that the project contributes to production diversification while it does not influence total revenues from sales, household welfare and diet. Chapter 4 shows that similar incentivized mechanisms elicit similar decisions in terms of monetary risk aversion only if other risk-related attitudes are accounted for. Furthermore, it examines whether individuals' characteristics and a self-assessed measure of risk aversion relate to individuals' choices in lotteries. The findings suggest that there is some external validity of the two studied tasks as predictors of self-reported risk attitudes.

Keywords: Social Capital, Corruption, Rural Development, Risk Aversion

Questa tesi utilizza metodologie differenti al fine di esplorare argomenti generalmente ascritti all'economia dello sviluppo. Il primo capitolo discute la letteratura sul capitale sociale scomponendolo nel suo componente strutturale, le reti, e cognitivo, la fiducia. Ogni componente è a sua volta scomposto in diverse sotto-dimensioni una delle quali, il particolarismo, è utilizzato nel secondo capitolo, sia a livello teorico che empirico, come determinante di forme di corruzione collusiva. Come previsto dalla teoria, il particolarismo ha un effetto positivo e causale sulla probabilità di offrire una tangente. Il terzo capitolo valuta l'impatto di un progetto di estensione agricola realizzato in Etiopia, volto ad introdurre la coltivazione di nuovi prodotti ortofrutticoli insieme ad alcune tecniche e strumenti innovativi. Empiricamente si utilizzano gli strumenti della valutazione d'impatto combinando confronti tra villaggi, attraverso una stima difference-in-differences, con una comparazione all'interno del villaggio usando uno studio controllato randomizzato. I risultati indicano che il progetto ha contribuito alla diversificazione produttiva ma non ha influenzato i ricavi ottenuti dalla vendita dei prodotti ortofrutticoli e, di conseguenza, il benessere delle famiglie. Il quarto capitolo mostra come meccanismi incentivati sufficientemente simili elicitino decisioni correlate in termini di avversione al rischio solo quando si tengono in considerazione altri atteggiamenti relativi al rischio. Inoltre si studia la correlazione tra l'avversione al rischio riportata e l'avversione al rischio ottenuta tramite lotterie. I risultati suggeriscono una misurata validità esterna dei due metodi studiati.

Parole Chiave: Capitale Sociale, Corruzione, Sviluppo Rurale, Avversione al Rischio

Contents

CHAPTER 1

The Metonymy of Social Capital
Why we should not take a part for the whole[1]

[1]The author wishes to thank Fabio Sabatini and Luca Stanca for useful comments and inputs.

1.1 Introduction

In recent years academic interest in the concept of social capital has grown fast. Inherited from sociology, the concept has been used in several disciplines, including political science, economics, education and anthropology. Nowadays, the term describes more a strand of the literature than a specific concept.[2] While this paved the way to a genuine interchange among scholars from different disciplines, the array of definitions and measurement methods used in the empirical literature has often made it difficult to compare the results of different studies and to formulate any general assessment about the effects of social capital (Sobel, 2002).

One of the reasons behind this difficulty is the practice, very common in economics, to use the label "social capital" to indicate one of its components, thus measuring a part for the whole. Such *metonymy of social capital"* is strictly related to the difficulty of measuring such a multifaceted phenomenon and has often undermined the credibility of the whole area of research (Fine, 2001). Nevertheless, the fact that social capital is commonly considered a key factor for economic development (Adhikari and Goldey, 2010), and that it has been associated with several socio-economic outcomes[3] make us believe that its importance can not be refused due to a lack of methodological clarity.

This survey defines social capital first and foremost as a micro phenomenon. Our goal is to show that by disentangling different components of social capital we are able to reason more effectively about its positive and negative effects and about their implications for economic and social development. First, we give a definition of social capital. Second, we present the two main components of social capital, networks and trust, respectively, by further decomposing them into their sub-dimensions. Trust, the *cognitive*

[2]For a discussion regarding whether the concept of social capital is indeed a good social science concept see Bjørnskov and Sønderskov (2013).

[3]Such as economic growth and poverty (Knack and Keefer, 1997), human capital (Coleman, 1988), development (Guiso et al., 2008), subjective health (Yip et al., 2007) and well-being (Sarracino, 2010), political accountability (Nannicini et al., 2013), the size of the shadow economy (D'Hernoncourt and Méon, 2012), innovative outcomes (Akçomak and Ter Weel, 2009), entrepreneurship (Bauernschuster et al., 2010), suicide (Helliwell, 2007), homicide (Rosenfeld et al., 2001) and crime rates (Buonanno et al., 2009)

component of social capital, is decomposed into general, particular and institutional trust. Networks, the *structural* component of social capital, are decomposed into three categories: bonding, bridging and linking. For each sub-component, we present how it has been measured in the empirical or experimental literature, make some critical points and highlight some open areas for future research. Third, we conclude by presenting a simple theorization of the micro-foundations of social capital. We show, in particular, how different types of trust are interrelated and how different kinds of networks produce different types of trust, thus producing negative or positive social capital.

1.2 Defining Social Capital

Social capital is generally referred to as those features of social life enhancing coordination and cooperation among people who pursue the same goal (Fukuyama, 1995; Putnam, 2001). Social capital arises from social networks and it is the use that individuals make of them that may produce social capital. As in Bourdieu (1986) and Coleman (1994), social capital is therefore *intangible*. In order to possess social capital, a person must be related to others and it is those others, not himself, who are the actual source of his social capital (Lin et al., 2001). To put it differently, for social capital to arise the existence of a social network is a necessary but not a sufficient condition. In fact, the existence of social capital depends on the quality of the networks, the action undertaken by individuals in building trust and reciprocity inside and towards those networks, and the resources available to their connections (Portes, 2000). Social networks and trust are valuable assets that enable people to build communities, to commit themselves to each other and, finally, to cooperate. This cooperation, however, brings some *benefits* and some *costs* by increasing the within-group welfare while decreasing the welfare of non members. These opposite effects are usually referred to as the *positive* and the *negative* side of social capital, and have been acknowledged in the literature for a long time.

The literature usually defines trust as the *cognitive* component of social

capital, while networks are generally referred to as its *structural* component (Burt, 2000). While trust is more linked to individuals' perceptions, and it is therefore more difficult to measure, networks are usually identified through observation of reality (e.g., participation in voluntary activities). The structural and cognitive components of social capital are either positively or negatively related (Sabatini, 2009b). Trust, for instance, can confer legitimacy to cooperative behaviors that can result in the formation of networks. These networks, in turn, strengthen trust and reciprocity. Conversely, certain types of networks hamper trust by restricting others, outside the network, in accessing it (Woolcock, 2001). Yet, the structural and the cognitive components of social capital can be further decomposed in sub-components, each one affecting several social and economic dimensions in a specific way. This approach to the study of social capital is very common in other social sciences but has not yet been extensively applied in economics.

1.3 Networks

Networks cover a wide terrain spanning from tight groups, such as the nuclear family, to more extensive ones, such as volunteer organizations. They are often themselves connected to one another, and accessing one of them could sometimes be costly for individuals. Networks are generally classified into bonding, bridging, and linking. Bonding social networks link homogeneous groups of people, bridging networks connect people belonging to different socio-economic groups, while linking networks refer to vertical connections with people in power.

The literature generally finds that bridging networks of individuals generate positive social capital, since they foster the diffusion of information and trust enhancing community cohesion. On the contrary, bonding networks are thought of as negative social capital burdens on society, since they foster societal interests at the expenses of the community.[4] It is therefore impor-

[4]Burt (1982) emphasizes the importance of the way different parts of networks are bridged. In particular, he points out the strategic advantage that may be enjoyed by individuals with ties into multiple networks that are largely separated from one another.

tant, in empirical works, to correctly disentangle the effects of bridging from those of bonding and linking networks. Yet, only a few scholars have thus far attempted to do so. Putnam (2001), for example, points out the importance of the distinction between bonding and bridging networks, but then does not really address the issue and does not perform separate analyses (Hibbing and Theiss-Morse, 2002). A (pioneering) example of the different effect that different kinds of social networks may have on - for example - development, is given in Banfield (1958) who partly attributed the backwardness of Southern Italy (compared to Northern) to the inability of citizens to

> act together for their common good or, indeed, for any end transcending the immediate, material interest of the nuclear family (p. 10)

A growing body of the literature suggests that networks are extensively used in the labor market (Ioannides and Loury, 2004; Pellizzari, 2010) both on the supply side, where social networks help workers to find a job, and on the demand side, where firms use networks to find better workers to fill vacancies (solving or trying to solve problems of adverse selection and moral hazard). Actually, the (intuitive) fact that information about jobs is passed through social networks becomes interesting when analyzing its implications for wage and employment dynamics. Some authors support the idea that the use of social networks may negatively affect behavior in markets. In fact, social networks can be used by low skilled but *connected* workers to be selected in place of more talented but unconnected individuals (i.e., discrimination in favor, nepotism, amoral familism). From another perspective, a labor market with referrals would improve the quality of job matches by reducing problems of adverse selection and moral hazard.

To date, the empirical evidence provides conflicting results. While some studies show that workers recruited through informal networks receive lower wages (Pellizzari, 2010), others find evidence that workers who have found their job through a social connection receive higher wages (Amuedo-Dorantes and Mundra, 2007). The literature seems to suggest that when studying the effects of specific networks in the labor market, the specificity of each network

translates into a peculiar effect on the outcome of interest. More generally, the literature shows that the effects of *professional* contacts (referrals from current employees, i.e. weak ties and bridging networks) have a positive effect on the quality of matches (in terms of wage, tenure, productivity, etc.) while *family* contacts (i.e., bonding networks) have a negative one.

Recently, some scholars have argued that referral hiring could be a cause of the persistence of inequality for labor force minorities (Calvo-Armengol and Jackson, 2004). This strand of the literature focuses on the negative role of social networks, leading to potentially self-perpetuating poverty traps (Durlauf, 2002). Beaman and Magruder (2012) deal in general with the problem of the activation of social networks in temporary and low-skilled labor markets. The authors create short-term jobs in the field in urban India by asking people drawn from a pool of active labor market participants to participate in a paid laboratory experiment and then to refer a friend or a relative to participate in it too. While everyone is asked to refer a friend who will be highly skilled for the job (i.e., who can replicate the task the referral did in the previous day), the type of referral contract and the amount offered is randomized: some are proposed a fixed amount while others are offered a guarantee plus a contingent bonus based on the referrals' performance. The results show that when individuals receive performance pay they become 7 percentage points less likely to refer relatives. This is a large change since less than 15 per cent of individuals refer relatives. They are also 8 percentage points more likely to refer co-workers.

1.3.1 Bonding Networks

Bonding networks refer to relationships between people who know each other and have several interactions among them, i.e., family members, kin, members of the same ethnicity or language group (Gittell and Vidal, 1998), and neighbors (DiPasquale and Glaeser, 1999). These relationships are defined "strong ties"by Granovetter (1973) and are considered the foundations on which other relationships with broader social networks are built. Generally, belonging to these networks is measured by considering the importance that

respondents assign to the family or to their closest group.

The empirical literature has often shown that bonding networks have negative effects on several economic outcomes. As an example, a study on Mexican immigrants by Amuedo-Dorantes and Mundra (2007) finds that social networks, particularly strong ties, contribute to the economic assimilation of immigrants by raising their hourly wages. However, strong ties allow for a lower employment likelihood. Sabatini (2009) examines the impact of four types of aggregate level social capital (strong family ties, networks with kin and friends, voluntary group affiliations and activities engagement, and political participation) on human development and well-being, finding that strong family ties and networks with kin and friends actually have negative effects on human development, although they improve life quality by reducing worker's precariousness.

Family

Family ties are the most commonly studied bonding network, since they often work in a peculiar way by changing individual behaviors and shaping economic institutions (La Porta et al., 1999). On the one hand, it is clearly recognized in the literature that the family is an institution that *substitutes* the state by internalizing social risk, by pooling resources across generations or by providing public goods and welfare services. On the other hand, since Weber (1904), the literature has often focused on the family as a potential form of bonding social capital. Coleman (1994), for instance, noticed that societies based on strong family ties promote codes of conduct according to which a selfless behavior is required within the small circle of kin whereas selfish behavior is acceptable outside this network. In the empirical literature, there is evidence showing that several economic behaviors and outcomes (individual and aggregate) are affected both by the structure of the family and by the strength of family ties.

The structure of the family is often classified along two dimensions, based on Todd's classification (Todd and Garrioch, 1985): vertical relationships between parents and children and horizontal relationships between siblings.

The first dimension is said to be liberal if children become independent of their parents at an early age and leave their parental home as soon as they get married, or authoritarian if children continue to depend on their parents in adulthood and still live with them after marrying. The second dimension defines the relationship as egalitarian if siblings receive an equal share of family wealth after their parents' death, or non-egalitarian when parents can favor one offspring at the expense of the others. Based on these two dimensions, four possible family types are identified: the absolute nuclear family, the egalitarian nuclear family, the stem family and the communitarian family.

Todd's classification has been used in several economic papers. Greif and Tabellini (2015) show that the presence of the nuclear family in Europe as opposed to the clan in China was central in explaining patterns of urbanization in Europe. Moreover, they associate generalized morality[5] with the diffusion of the nuclear family in Europe, whereas limited morality was typical of China, where the clan was more diffused. Duranton et al. (2009) use Todd's classification to explain regional disparities across Europe in household size, educational attainment, social capital, labor force participation, sectoral structure, wealth, and inequality. Galasso and Profeta (2013) use Todd's classification to show that family structures are crucial for explaining different types of pension systems and that Todd's definition of nuclear and extended family is strongly correlated to the measure of family ties as defined in Alesina and Giuliano (2011). The literature on family structure has also examined, among other things, the relationship between family structure, inheritance norms and the performance of family businesses (Pérez-González, 2006), children's achievements (Ermisch and Francesconi, 2001; Gennetian, 2005) and child poverty (Lerman, 1996), cooperation (Greif and Tabellini, 2010) and farmers' decisions to adopt a new crop (Bandiera and Rasul, 2006).

Recently, scholars have focused, on the strength of family ties, rather than the structure. Bentolila and Ichino (2008) study how countries with different family ties cope with unemployment shocks. They find that the consumption

[5]The distinction between "generalized" versus "limited" morality was first introduced by Platteau (2000) and stressed by Tabellini (2010).

losses after the termination of a job are much lower in Mediterranean Europe, due to strong family ties. Ermisch and Gambetta (2010) find that people with strong family ties put a lower level of trust in strangers than people with weak family ties. Alesina and Giuliano (2010) study the importance of family ties for economic behavior. Using data from the World Values Survey (WVS), they show that with strong family ties home production is higher and families larger, whereas labor force participation of women and youngsters and geographical mobility are lower. They address causality by looking at the behavior of second generation immigrants. Alesina and Giuliano (2011) show that there exists an inverse relationship between family ties and political participation. In particular, the more individuals rely on the family as a provider of services, insurance, and transfer of resources, the lower is one's civic engagement and political participation. Moreover, they show that strong family ties appear to be a substitute for generalized trust, rather than a complement. Alesina et al. (2015) study the relationship between strong family ties and the regulation of labor. They find that individuals who inherit stronger family ties are less mobile, have lower wages, are less often employed and support more stringent labor market regulations. At the aggregate level, they find a positive cross-country correlation between the strength of family ties and labor market rigidities. On the contrary, a positive correlation is found between labor market rigidities at the beginning of the 21st century and family values prevailing before World War II, which suggests that labor market regulations have deep cultural roots. On the positive side, Alesina and Giuliano (2014) find that family relationships are found to improve well-being, as measured by self-reported indicators of happiness and subjective health. At the aggregate level, the authors find a negative correlation between the strength of family ties, economic development and the quality of institutions measured by the Worldwide Governance Indicators (WGI) of the World Bank. Although they do not offer any definite answer to the question of causality, they show that family values are quite stable over time and that they determine institutional differences and differential levels of development across countries.

Kins

A peculiar aspect linked to the family is the so-called *kin system*. Following Wolf (1955), this can be defined as *a system of shared rights and obligations encompassing a large number of near and distant relatives.* The kin system is a social contract of mutual assistance among members of an extended family that allows people to reduce their exposure to risk and to cope with financial markets failures (Fafchamps and Lund, 2003; Jackson et al., 2012). As reported in Di Falco and Bulte (2011), kinship is the fundamental social organization in less developed countries, regulating social relationships, social customs and the access to resources and services. Unlike other types of networks, membership in the kin is not voluntary and it implies moral obligations for its members, such as sharing and redistribution. These obligations are supported by customs and norms.

As formally explored by Hoff and Sen (2005), during economic transition the kin system can become a poverty trap for its members since it imposes barriers in entering the market economy such as workplace nepotism or inefficient transfers to kin. To date, the actual impact of the kin system on economic development is unclear[6] and the consequences of membership in a kin group vary for people in different parts of the world. As an example, while in African firms there is little evidence of positive effects of the kin system on productive activities, the kin system has spurred entrepreneurship in East Asia (Di Falco and Bulte, 2011).

Some recent experimental evidence shows how differences in kinship determine the choices made by subjects in trust and ultimatum games. Barr (2004), for instance, makes a behavioral comparison between two groups in Zimbabwe: a group of new villages that were set up in 1997 as resettlements consisting almost entirely of unrelated households, and a control group of non-resettled villages made up almost exclusively of kin. She finds no differences in socially transmitted behavioral rules, in altruism and in loyalty between resettled and non-resettled villages. However, she finds that reset-

[6]For a more detailed discussion of the relationship between kinship groups and economic outcomes see La Ferrara and Milazzo (2014) while for a review of the relevance of kinship ties in development see La Ferrara (2007).

tled villagers have lower level of trust than non-resettled villagers, a symptom of a lower density in kinship ties, which in turn leads to less familiarity with other people's behavioral characteristics.

1.3.2 Bridging Networks

Bridging networks are horizontal ties connecting heterogeneous groups of people. In the empirical literature, they have been proxied, among other things, by membership in any kind of association, blood donations and by the number of readers of local newspapers. The choice of mono-dimensional indicators is not free from criticism. First, as suggested by Sabatini (2009a), the use of mono-dimensional indicators generates confusion about the relationship between social capital and its effects, since social capital is actually measured by its effects. The tautology, in this case, lies in the attempt to measure a phenomenon by an indicator of its outcome. As an example, when social capital is measured as the percentage of blood donations (Nannicini et al., 2013), it is necessarily proxied by a (positive) outcome of its presence. This tautologically results in a positive effect of social capital on any social and economic phenomenon. Second, this strategy assumes all types of associations to be equally bridging. However, most groups are likely to be both bridging and bonding to some extent. As an example, religious organizations often have hierarchical structures which may negatively affect the cooperative attitudes of members (van Deth, 2010) while cultural and educational associations may be conducive to particular interests which do not necessarily entail relational activities or pro-social motivations and behaviors (Degli Antoni, 2009). Hooghe (2003), for instance, argues that voluntary associations contribute to strengthening the values that dominate a certain group. As an example, let us suppose that anti-racism is considered as a positive value for a society. If a group or an association is mainly composed by highly educated members who exert a preference for anti-racist sentiments, then group interactions strengthen these anti-racist values (Stolle and Lewis, 2002). On the contrary, if a group is mainly composed by members with ethno-centric attitudes, then group interactions strengthen this ethno-centrism. Thus, the

bridging degree of any association is inevitably affected by its members' heterogeneity, by their degree of involvement into associational life and by the relative size of each group (Blau, 1977).

To address the problem of the significant diversity among associations, Paxton (2002) proposes to make use of the interconnections between voluntary associations and the diversity of their membership. By doing so, they are able, to some extent, to divide associations according to their bridging potential. Grießhaber and Geys (2012), find that the impact of membership on corruption varies significantly according to the association's characteristics in terms of inclusiveness and interconnectedness.

1.3.3 Linking Networks

Linking networks are ties connecting individuals to people or groups in powerful positions. A leading example of these networks are civil society organizations that allow citizens to interact with the institutions to carry out advocacy activities through, for instance, collective action. These networks are usually thought to produce civic capital. However, as underlined in Knack and Keefer (1997):

> if the economic goals of a group conflict with those of other groups or of unorganized interests, the overall effect of group memberships and activities on economic performance could be negative. Although the ability of groups to articulate their interests is likely to be an important restraint on government, it also provides groups a way to capture private benefits at the expense of society (p. 1271)

As a result, linking networks may have positive as much as negative effects and, as for bridging networks, they should be operationalized in a multidimensional way. Despite this, previous studies have often proxied linking networks by indexes measuring participation in political activities. These indexes are normally composed of very different activities, such as participating in political meetings, boycotting products or wearing a sticker, the

percentage of civically-oriented establishments in a community, voting, and indicators of philanthropic activities such as charity and leisure. Given the lack of unambiguous conceptualizations and explanations, studies relying on a multitude of measures of linking networks have often generated equivocal findings and interpretations (Durlauf, 2002).

Three main problems have been raised in the literature on linking networks. First, it is necessary to analyze separately the structure from the level of participation. Ekman and Amnå (2012), for instance, propose to consider political participation to be composed of a manifest and a latent part. The manifest political participation comprises parliamentary and extra-parliamentary forms of activism while the latent component refers to activities based on personal interest in politics and societal issues. In the latter form of political participation, the extrinsic motivations are sharply prevailing, at the point that these forms of political participation are actually rent seeking, and maintaining connections with people in power becomes crucial to spur personal (and in many cases antisocial) interests. This is the case of Mafia-type organizations (Çayli, 2010). Second, the indicators proposed above often include both participation and actions, while quite often participation does not immediately translate in concrete actions. Third, there exists an important problem related to causality. In fact, several studies assume that social networks mobilize people into political engagement. However, it may also be the case that political participation in itself affects the types of networks one belong to thus suggesting an opposite causal flow (Quintelier et al., 2011).

1.4 Trust

Social scientists view trust as an economic lubricant that facilitates cooperation, thus fostering economic activity. In general terms, trust is the decision to rely on others (i.e., a person, a group, a firm, an institution) in a risky situation. According to Arrow (1972)

> virtually every commercial transaction has within itself an el-

ement of trust, certainly any transaction conducted over a period of time. It can be plausibly argued that much of the economic backwardness in the world can be explained by the lack of mutual confidence. (p.357)

For Fukuyama (1995) trust is:

> both the condition for, and the effect of, the forms of social capital - collective values, social networks and cultural mores - that underpin social cohesion and shape economic growth.

The sociological literature decomposes trust in several sub-dimensions. If trust occurs towards unknown members of society, then it is defined as *general trust*. General trust is context-dependent, influenced by experience and, at the society level, it is affected by some shared ethical and moral norms of reciprocity (Fukuyama, 2001). General trust differs from institutional trust. If the first one is generally *horizontal*, the second one is *vertical*, since it includes trust in social system (Hayoz and Sergeyev, 2003), in institutions (both formal and informal ones) and in their officers (Hardin, 2002). General trust also differs from particular trust. While the former concerns unknown others, the latter involves a narrow circle of familiar others. Uslaner (2002) defines *general trust* as *moral trust*. Moral trust is a stable trait which exists regardless of the context, regardless of prior experiences and regardless of the recipient. The moral dimension of trust comes from the fact that trusting strangers requires accepting them into our moral community without having any personal knowledge of them but assuming that they all share some fundamental moral values. Moral trust leads people to treat others as if they were trustworthy (Uslaner, 2008b).[7]

1.4.1 General Trust

Within economics the most widely employed measure of generalized trust is the question, introduced by Rosenberg (1956) and used cross-nationally

[7]For a discussion of the micro and macro foundations of trust see Uslaner (2008a)

in the World Values Survey (WVS), the General Social Survey (GSS), the European Social Survey, and most of the Barometers (the Latino Barometer, the Afrobarometer, the Asian Barometer, etc.): "Generally speaking, would you say that most people can be trusted or that you can't be too careful when dealing with others?". Possible answers are typically either, "Most people can be trusted" or "Need to be very careful".[8]

A large number of papers have analyzed the role of trust in economics. This extensive literature, recently reviewed by Algan and Cahuc (2013), has been challenged on several grounds. The first set of criticisms concern the trust question itself. First, it is argued that the respondents are given only the possibility to choose between trust and caution, rather then between trust and distrust, whereas trust and caution are not always mutually exclusive (Naef and Schupp, 2009). As an example, Miller and Mitamura (2003) have shown, that Japanese students are more trusting than American ones. However, when measuring trust and caution separately, they find that Americans are more trusting than Japaneses and, at the same time, also more cautious. Second, the answers to the trust question may differ significantly depending on whether people understand *most people* in the question as meaning acquaintances or strangers (Reeskens and Hooghe, 2008). Recently, Delhey et al. (2011) have addressed this question, finding that in 41 out of 51 countries, *most people* in the standard question predominantly connotes out-groups. To this extent, it is a valid measure of general trust in others. Nevertheless, the radius of *most people* varies considerably across countries. Some country rankings for trust change when the standard question is replaced by a radius-adjusted trust score. Since the assumed civic nature of trust depends on a wide radius, then the radius of trust matters for civic attitudes and behaviors.

Another set of criticisms regards the fact that the answer about trust is always context-dependent (Goudge and Gilson, 2005) and that this intrinsic

[8]In some other papers trust is elicited by constructing an index using either the "fairness question" contained in the WVS ("Do you think most people would try to take advantage of you if they got the chance, or would they try to be fair?") or the "help question" contained in the GSS ("Would you say that most of the time people try to be helpful, or that they are mostly just looking out for themselves?")

characteristic leads to biased empirical results when passing from individual to aggregate variables. The macro-level studies on social capital generally use a trust indicator that is given by the percentage of respondents replying "most people can be trusted" to the Rosenberg question. Trust represents individuals' perception of their social environment and is, therefore, inevitably related to the position people occupy in the social structure. By aggregating individual data, every linkage with the social and historical circumstances in which trust and social capital are located are inevitably lost, therefore altering the concept of trust (Foley and Edwards, 1999). According to this criticism, aggregate levels of trust are, in fact, a proxy for the well-functioning of institutions (Beugelsdijk, 2006). In a similar vein, Bjørnskov (2006) suggests that social trust has a positive impact on the quality of government, whereas Rothstein (2005) argues that it is rather good government that causes general trust. Uslaner (2008a) strongly refuses this critique on the basis of the difference between what he calls experience-based trust and moralistic trust. In his view, while experience-based trust is fragile, and is based upon responses to how others have previously treated you and to your interactions with others in a specific context, moralistic trust is not. According to Uslaner (2008a) the academic literature based on the Rosenberg's question reflects, in fact, moralistic trust and not experience-based trust. However, Ahlerup et al. (2009) show that the effect of social capital on economic growth, as well as on the investment rate, is nonlinear and dependent on the quality of formal institutions. More specifically, social capital matters the most when formal institutions are weak and almost ceases to matter when institutions are strong.

Recently, a stream of research has combined survey and experimental methods to test the relationship between behavioral and attitudinal measures of trust. In contrast to survey measures, experimental measures of trust between anonymous individuals and with monetary incentives can capture generalized trust behavior more effectively (Camerer, 2010) since they allow to use actual choices as an appropriate way to infer preferences. However, behavioral measures of trust can potentially suffer significant shortcomings. In fact, behavioral measures of trust are primarily elicited through

the trust game (Berg et al., 1995).[9] However, some papers have recently shown that behavior in the trust game often correlates with altruism (Cox, 2004; Schechter, 2007), thus rising the question of the accuracy of behavioral methods, such as the trust game, in measuring trust (McEvily et al., 2012).

The empirical evidence on the correlation between the trust question and experimental measures of trust is mixed. Glaeser et al. (2000) find that standard attitudinal survey questions about trust predict trustworthy behavior in the experiments much better than they predict trusting behavior. To put it otherwise, trust as elicited through the standard trust question has no predictive power for trust as measured in the trust game, while it predicts trustworthiness. In contrast, Fehr et al. (2003) find a relation between survey-measured trust and experimentally measured trust, but not trustworthiness, in a representative sample of the German population. Sapienza et al. (2013) find that the World Values Survey (WVS) question on trust captures mostly the belief-based component of a trust game. Yet, Bellemare and Kröger (2007) find no relation at all neither with trust nor with trustworthiness. Holm and Danielson (2005) find a strong and significant relation between Dictator donations and proportions returned in Trust games among undergraduates in Tanzania and Sweden. They find that, although using the same protocol, the predictive power of survey trust questions differed between countries. While they found a relation between survey trust and trust behavior in Sweden, they found no relation at all in Tanzania. Johnson and Mislin (2012) use a constructed data set containing observations of trust and trustworthiness behavior from replications of Berg et al. (1995) collected across 35 countries from more than 23,000 subjects, finding strong correlation with the trust survey question. By contrast, they find no relationship

[9]Readers interested in a detailed review of results for the Trust Game are referred to Cardenas and Carpenter (2008) and Johnson and Mislin (2011). In this game two are players endowed with some fixed amount of money. The first player, the sender, moves first by choosing how much money to send to the second mover, the receiver. The amount sent is increased by the experimenter according to a commonly known function before being given to the second player. The second player is then asked to give some amount back to the first player, ending the game. Under standard economic assumptions, the subgame perfect Nash equilibrium of the one-shot game is that the receiver never returns any money and that, in turn, the sender never sends any money. Trust in others is measured through the sender's behavior, whereas trustworthiness through the receiver's behavior.

between trustworthiness and the trust question.[10]

1.4.2 Particular Trust

In contrast with the literature at the macro level, numerous types of trust have been developed at the micro level. In economics, a growing attention has been devoted to the fact that trust can be classified in two categories: in-group and out-group trust. If out-group trust involves confidence in strangers, and is in fact equivalent to general trust (Delhey et al., 2011), in-group trust concerns confidence in people who share the same values and who are connected by social ties or social identities (Uslaner, 2002). Another way to define the same categories is to use the terms *particular* and *general* trust, *particularized* and *generalized* trust (Uslaner and Conley, 2003), or *thick* and *thin* trust (Newton, 1997).

Particular trust arises in face-to-face interactions and can be thought of as reputation (Bjørnskov, 2007). According to Uslaner (2002),

> The difference between generalized and particularized trust is similar to the distinction Putnam (1993) drew between bonding and bridging social capital. We bond with our friends and people like ourselves. We form bridges with people who are different from ourselves. While Putnam argued that both can lead to trust, he held that bridging organizations would produce much more trust. (p. 7)

While confidence in unknown others is widely recognized as an essential component of positive social capital, since it leads to widespread cooperation (Putnam, 2001), confidence only in known people is usually conducive to negative (or anti-social) capital (Banfield, 1958; Foley and Edwards, 1999). In fact, as underlined in Baron (2010), agents overcome the incentive to free-ride on other people when they interact with people who are close to them.

[10] Another strand of the literature looks at the relationship between trust experimentally elicited and cooperation. For a review of the literature see Thöni et al. (2012)

At present, the study of particular trust has not met with great success in economics. Yet, there are many possible and interesting applications. If particularized trust is conducive to negative social capital, then it can constitute an impediment to economic development and democratization (Bo Rothstein, 2008). Indeed, it could also become a self-feeding trap in the sense that when particular trust becomes a key feature of state power, particular trust itself is generalized and general institutions become the way to perpetuate particular interests. This is the case, for instance, of the Mafia-based societies (Gambetta, 1996). Moreover, as suggested by Fukuyama (1995), if people only trust those with whom they interact regularly and if trust decreases with social distance, then the extent of this reduction determines the size of the market.[11] On the opposite, particular trust may translate into general trust. This is what Putnam (1995) defines the transitivity of trust. One of the possible mechanisms that guarantees this transformation is presented in Macy and Skvoretz (1998). In their framework, cooperation can emerge between strangers even without formal or informal social controls. In particular, general trust emerges in local settings in exchanges between neighbors and spread through weak ties to outsiders.

With exception of Guiso et al. (2009),[12] showing that In Europe managers from any given country express more trust in people from that country than managers from other European countries, the literature on particularized trust is mainly experimental and is based on the analysis of individual-level behavior using the trust game when group identity is hidden and when it is not. Much of this literature has found that people transfer more money to known than unknown others.

In Fershtman and Gneezy (2001) and Fershtman et al. (2005) subjects

[11] van Rijn et al. (2012) show, for instance, that structural social capital, especially in the form of connections beyond the village, is associated with more extensive adoption of innovations, the reverse is true for cognitive social capital measured as trust within the local community.

[12] In a recent paper using the Bank of Italy's Survey on Household Income and Wealth (SHIW), Albanese et al. (2013) adopt a behavioral approach and distinguish between general trust and trust in known people. They show that time preferences and risk preferences are key covariates of self-reported trust and that both predict negatively a measure of generalized trust. On the opposite, risk aversion is positively correlated with an index of particularized trust (which refers to family and friends).

play the Trust Game with different ethnic groups. Fershtman and Gneezy (2001) find a systematic mistrust toward men of Eastern origin while a dictator game experiment indicated that this discrimination was due to (mistaken) ethnic stereotypes and not to a taste for discrimination. Fershtman et al. (2005) employ the trust game to compare the amounts sent by player A to individuals with clear group affiliation and individuals with no clear group affiliation (i.e. anonymous individuals) to determine whether the player A's behavior can be characterized as nepotism or discrimination. More specifically, they use this experimental test to analyze linguistic segmentation between the Flemish and Walloon in Belgium, and religious segmentation between ultra-orthodox and secular Jews in Israel.

Brandts and Solà (2010), implement the trust game between friends and others, showing that friends are more likely to send each other higher amounts, even under the presence of social inefficiency when the productivity of in-group members is varied. A similar result is found in Brañas-Garza et al. (2012). Buchan et al. (2006), randomly assign people to artificially created groups. Although this study involves students from different countries, students are paired only with participants from their own country. Before playing the Trust Game, participants spend ten minutes in groups. When the Trust Game is played each participant know whether he is playing against someone from their discussion group (the in-group) or someone from a different group (the out-group). They find heterogeneous effects of this experimentally group membership across countries. While Americans send (and return) more money to in-group members, Chinese students send (and return) more to out-group members.

Buchan and Croson (2004) examine the boundaries of trust and trustworthiness in the United States and China. More specifically, participants play a standard Trust Game against someone from their own university. After the Senders have made their transfers, but before they find out how much he returns back, subjects are asked to fill a questionnaire containing a hypothetical question regarding how much money they would send (or return if they are a Recipient in the game) if they were to play the game against seven different types of player with varying degrees of social distance: a parent, a

sibling, a cousin, a student they know well, a student from another university, a stranger from their home town and a stranger from another country. They find that, both in the United States and in China, the hypothetical amounts sent and returned tend to fall as social distance increases. Etang et al. (2011) use experimental data, elicited through the trust game, from rural Cameroon to quantify the effect of social distance on trust and altruism. Subjects are asked to play a Trust Game with fellow villagers or with someone from a different village. They find that significantly more money is sent when the players are from the same village. Other factors that influence transfers at least as much as the same-village effect are gender, education and membership of rotating credit groups. Moreover, to test whether senders are motivated by altruism, they ask subjects to play a Triple Dictator Game. Senders transfer significantly more money on average in the Trust Game than in the Triple Dictator Game. However, there is also a social distance effect in the Triple Dictator Game. Furthermore, results from a Risk Game suggest that Trust Game transfers are uncorrelated with attitudes towards risk.

Vivian Lei (2010) studies how income inequality activate in-group favoritism and, if so, whether such a bias is strong enough to survive the removal of income inequality. The results suggest that there exists in-group favoritism only on the part of rich first movers. Rich first movers trust their in-group members significantly more in the presence of income inequality. Poor first movers, in contrast, do not exhibit such in-group bias. They do not discriminate between in-group and out-group. Moreover, in-group and out-group favoritism established in the past can be alleviated, but not completely removed, by an equal income distribution.

These results, although interesting, leave open a large hole in the empirical literature. On the one hand: What are the individual-level determinants of particular trust? Which kind of networks generate particular or general trust?[13] On the other hand: Do the real data confirm the theoretical relationships between particularized trust, negative social capital and development?

[13]A first attempt to answer to this question is given in Iglič (2010).

1.4.3 Institutional Trust

Institutional trust captures the degree of trust that people place in different institutions. Differently from general and particular trust, institutional trust focuses therefore on an impersonal object. Institutional trust can be defined as the security one feels about a situation because of guarantees, safety nets, or other structures (McKnight et al., 1998). Institutions can be formal (i.e., the legal system, the educational and vocational system, the financial system, the government, the army etc.) or informal (i.e., customs). Informal practices become important institutions when they are stable and commonly acknowledged. Institutional trust is therefore closely linked to general trust and civic engagement. In a study by Pearce et al. (2000), a minimum level of institutional trust is regarded as a *sine qua non* for the emergence of interpersonal trust. Therefore, as suggested by Fukuyama (2001), social capital and the institutional environment are, in a sense, complementary. Institutional trust links ordinary citizens to the institutions that are intended to represent them, thereby conferring them legitimacy. At the same time, the emergence of interpersonal trust depends on the extent to which institutions provide an environment in which trust can be rewarded and not exploited (Knack and Keefer, 1997).

Empirically, institutional trust is either measured directly in terms of whether people trust the government and politicians to do the right thing, or elicited through the answer to some specific question asking the respondents to report their degree of trust from 1 to 10 in several institutions.[14] A survey of the literature on institutional trust shows that most studies deal with formal institutions (Tan and Tambyah, 2011). As an example, institutional trust has been shown to have an effect on the support for government spending, re-distributive policies (Rudolph and Evans, 2005), compliance (Torgler, 2006), pro-social behaviors (Andriani and Sabatini, 2013), and on and support for law compliance (Marien and Hooghe, 2011).

Although it is important to examine trust in formal institutions, informal

[14]See for instance the WVS, the ESS and most of the Barometers (the Latino Barometer, the Afrobarometer, the Asian Barometer, etc.)

or non-political institutions, and the relationship between interpersonal trust and trust in these institutions, may play a key role for development (Bratton, 2007). If trust confers legitimacy to institutions, therefore perpetuating their role, then informal-institutional trust may play a crucial role. To the best of our knowledge, thus far, no studies have investigated this relationship given the lack of data.

1.5 Network(s) and Trust(s)

So far, we have seen that social capital, as an umbrella-definition, can be further decomposed in several sub-dimensions producing either positive or negative social capital. We have defined social capital as being composed by a structural and a cognitive component underlying that both components are inextricably linked to each another. The question now is to understand how the sub-dimensions of each component are linked. In particular, we ask two questions. First, how general trust is related with particular and institutional trust? Second, which kind of networks foster the diffusion which kind of trust?

The experimental and the theoretical literature suggest that general trust is associated with beliefs in other people's trustworthiness (Sapienza et al., 2013), the preferences of the recipient (Thöni et al., 2012) and that people rely on preexistent cognitive schema regarding the general cooperativeness of individuals and organizations (Robbins, 2016). Therefore, generalized trust, as well as particularized trust, are context and experience dependent. Indeed, it is the radius of experiences and predispositions that matters for the radius of trust (Freitag and Traunmüller, 2009). Generalized trust appears to have both cultural and institutional foundations. Dinesen (2013), by looking at immigrants having migrated from a broad range of countries of origin to destination countries in Western Europe, examines how their generalized trust is affected by the culture of their country of origin (in terms of the level of trust of this country) as well as institutional quality in the country they have migrated to (in terms of freedom from corruption). The results show that controlling for confounding variables, both factors have a highly

significant impact on trust. However, the relationship between institutional quality and generalized trust is reciprocal. Robbins (2012) uses data from the World Values Survey, World Bank, and other data sources in an identified non-recursive structural equation model, to show that generalized trust and institutional quality form a positive reciprocal relationship, where the connection is stronger from generalized trust to institutional quality.

Generalized trust is strongly related also with particularized and institutional trust (Freitag and Traunmüller, 2009). Newton and Zmerli (2011) show, both at the micro and macro level, that particularized trust seems to be a necessary but not sufficient condition for general trust, and both forms of trust appear to be necessary, but not sufficient conditions for institutional trust. At the same time, institutional trust is causally related with social trust. Sønderskov and Dinesen (2015) make use of two Danish panel surveys to address the potentially reverse and/or spurious relationship between institutional and generalized trust. Using individual fixed effects and cross-lagged panel models, they provide strong evidence of trust in state institutions exercising a causal impact on social trust, whereas the evidence for a reverse relationship is limited.

Generalized trust is related also to ethnic diversity, but the size of the effect is mediated by the level of socialization and by individual-attitudes towards diversity. Stolle and Harell (2013) build on the insights of the contact hypothesis and political socialization literature to go beyond recent findings that racial and ethnic diversity have overwhelmingly negative effects on social capital, particularly generalized trust. Using the Canadian General Social Survey (2003), the paper shows that, despite a negative relationship among adults, younger Canadians with racial and ethnic diversity in their social networks show higher levels of generalized trust. The results seem to confirm that youth socialization experiences with rising diversity and the normalization of diversity in a multicultural environment contribute to beneficial, rather than detrimental (Alesina and La Ferrara, 2002), effects of diverse social networks. Moreover, there is another form of diversity that can be fundamental in determining the levels of trust and social capital, namely value diversity. Beugelsdijk and Klasing (2015) focus on the extent

to which key human values and beliefs are shared in society, which captures a dimension of diversity not previously discussed. The paper finds that value diversity, in particular with regard to political ideological values concerning income redistribution and the role of the government in influencing markets, is important for understanding the international variation in trust, with high diversity being associated with lower levels of trust. This relationship is robust to controlling for various other determinants of trust, including other dimensions of diversity, and holds at various levels of aggregation. At the micro-level the opposite conclusion is reached by Quintelier et al. (2011). The paper shows that young people who have more diverse social networks are in fact more likely to participate in a variety of social and political activities. In particular, political diversity in one's social network has the strongest positive effect on political participation, but ethno-cultural diversity also matters. However, in terms of causality, they find a reciprocal cross-lagged effect, suggesting that political participation strengthens people's network diversity and diversity spurs political participation (controlling for prior participation and diversity).

Sabatini (2009b) studies the causal relationships between four types of social networks (i.e., bonding, bridging, linking, and corporate), and different forms of trust (knowledge-based trust, social trust, trust towards public services and political institutions) in a community of entrepreneurs located in an industrial district in Italy. The paper suggests that the main factor fostering the diffusion of generalized trust among entrepreneurs is the establishment of corporate ties through professional associations. Trust in people is positively and significantly related also to higher levels of satisfaction and confidence in public services, while participation in volunteering organizations does not appear to increase generalized trust, a result that has been recently confirmed by Bekkers (2012). Recently, Ermisch and Francesconi (2001) and Alesina and Giuliano (2011) have found that people with strong family ties have a lower level of trust in strangers than people with weak family ties

As already underlined above, in the social capital literature the micro-relationship between membership in voluntary associations and generalized

trust have been so far taken as an assumption (Rothstein and Stolle, 2008). As a consequence, we actually do not know exactly which kind of social interactions are sufficient and necessary for the institutionalization of cooperative behaviors and generalized trust and which kind are conducive, instead, to particularized trust. Iglič (2010) shows that when members of voluntary associations build particularized trust rather than generalized trust, this decreases their levels of social and political tolerance. Since associations are social contexts within which processes of interpersonal influence and political mobilization take place, they reinforce civic as well as un-civic orientations of associational members. Indeed, it is the scope (Wollebæk and Strømsnes, 2007; Isham et al., 2006) or the multiple affiliation to several associations with different purposes (Wollebaek and Selle, 2002), more than the activity level of members, that matters for the formation of social capital. However, a problem of selective attrition emerges. People self-select into association groups, depending on their original levels of generalized trust and reciprocity (Bekkers, 2012). The empirical literature must therefore carefully address the question of causality before claiming any effect.

Moreover, although in the economic literature intrinsic motivations have received much attention in explaining human behavior, quite surprisingly the existing literature on social capital fails to consider intrinsic motivations as a source of social capital. One relevant exception is Degli Antoni (2009). The paper studies the role that different motivations (ideal motivations, the desire to feel useful to others, the pursuit of social recognition and the desire to increase the number of acquaintances or friends) to volunteer have on the creation of volunteers' social capital, which is intended in this paper as networks of cooperative relations. The author finds that intrinsic motivations enable people to extend their social networks by creating relations characterized by a significant degree of familiarity. By contrast, extrinsic motivations, and in particular the decision to join an association in order to increase the number of acquaintances or friends, promote the creation of networks from a quantitative point of view, but they do not facilitate the creation of relations based on a particular degree of confidence.

1.6 Conclusions

This survey claims that the very common practice in economics to use the label "Social Capital" to indicate one of its components, measuring unidimensionally a multifaceted phenomenon, has undermined the credibility of the whole area of research. In order to regain credibility the empirical literature has to go back to the theory and, instead of focusing only on a component to measure social capital, should adopt a multidimensional prospective. An effort is therefore required in further decomposing both the structural and the cognitive components of social capital into their sub-components to study their effects on different economic dimensions. In fact, when disentangling different types of social capital's sub-components, the results may allow for a better understanding of the phenomenon observed greatly reducing the confusion that often surrounds the empirical research on social capital.

Despite the large number of papers produced on social capital, there are still many areas in which this concept could be usefully applied once correctly decomposed. The space to develop new research is very broad and the possibility of relevant policy implications is substantial, especially because changing habits regarding the use of networks and trust takes a long time and requires strong incentives.

I want to conclude this survey in a non-standard way, using Saint-Exupéry's words:

> La grandeur d'un métier est peut-être, avant tout, d'unir les hommes: il n'est qu'un luxe véritable, et c'est celui des relations humaines. (Terre des hommes, p.42, Livre de Poche)[15]

These words are, in themselves, a summary of the motivations of this survey and the reasons why it is still sensible to put effort in the study of the networks and trust linking individuals in societies.

[15]The size of a business is perhaps above all to connect men: it is a real luxury, and it is that of human relationships.

Bibliography

Adhikari, K. P. and Goldey, P. (2010). Social capital and its downside: the impact on sustainability of induced community-based organizations in nepal. *World Development*, 38(2):184–194.

Ahlerup, P., Olsson, O., and Yanagizawa, D. (2009). Social capital vs institutions in the growth process. *European Journal of Political Economy*, 25(1):1 – 14.

Akçomak, I. S. and Ter Weel, B. (2009). Social capital, innovation and growth: Evidence from europe. *European Economic Review*, 53(5):544–567.

Albanese, G., De Blasio, G., and Sestito, P. (2013). Trust and preferences: Evidence from italy's survey data. *Bank of Italy, mimeo*.

Alesina, A., Algan, Y., Cahuc, P., and Giuliano, P. (2015). Family values and the regulation of labor. *Journal of the European Economic Association*.

Alesina, A. and Giuliano, P. (2010). The power of the family. *Journal of Economic growth*, 15(2):93–125.

Alesina, A. and Giuliano, P. (2011). Family ties and political participation. *Journal of the European Economic Association*, 9(5):817–839.

Alesina, A. and Giuliano, P. (2014). Family ties. *Handbook of Economic Growth*, 2:177–215.

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

Algan, Y. and Cahuc, P. (2013). Trust, growth and well-being: New evidence and policy implications. *CEPR Discussion Paper No. DP9548*.

Amuedo-Dorantes, C. and Mundra, K. (2007). Social networks and their impact on the earnings of mexican migrants. *Demography*, 44(4):849–863.

Andriani, L. and Sabatini, F. (2013). Trust and prosocial behaviour in a process of state capacity building: the case of the palestinian territories. *Journal of Institutional Economics*, pages 1–24.

Arrow, K. J. (1972). Gifts and exchanges. *Philosophy & Public Affairs*, pages 343–362.

Bandiera, O. and Rasul, I. (2006). Social networks and technology adoption in northern mozambique*. *The Economic Journal*, 116(514):869–902.

Banfield, E. C. (1958). The moral basis of a backward society. *New York: Glencol*, 3:4.

Baron, D. P. (2010). Morally motivated self-regulation. *American Economic Review*, 100(4):1299–1329.

Barr, A. (2004). Forging effective new communities: The evolution of civil society in zimbabwean resettlement villages. *World Development*, 32(10):1753–1766.

Bauernschuster, S., Falck, O., and Heblich, S. (2010). Social capital access and entrepreneurship. *Journal of Economic Behavior & Organization*, 76(3):821 – 833.

Beaman, L. and Magruder, J. (2012). Who gets the job referral? evidence from a social networks experiment. *The American Economic Review*, pages 3574–3593.

Bekkers, R. (2012). Trust and volunteering: Selection or causation? evidence from a 4 year panel study. *Political Behavior*, 34(2):225–247.

Bellemare, C. and Kröger, S. (2007). On representative social capital. *European Economic Review*, 51(1):183–202.

Bentolila, S. and Ichino, A. (2008). Unemployment and consumption near and far away from the mediterranean. *Journal of Population Economics*, 21(2):255–280.

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

Beugelsdijk, S. (2006). A note on the theory and measurement of trust in explaining differences in economic growth. *Cambridge Journal of Economics*, 30(3):371–387.

Beugelsdijk, S. and Klasing, M. J. (2015). Diversity and trust: the role of shared values. *Journal of Comparative Economics*.

Bjørnskov, C. (2006). The multiple facets of social capital. *European journal of political economy*, 22(1):22–40.

Bjørnskov, C. (2007). Determinants of generalized trust: A cross-country comparison. *Public choice*, 130(1-2):1–21.

Bjørnskov, C. and Sønderskov, K. (2013). Is social capital a good concept? *Social Indicators Research*, 114(3):1225–1242.

Blau, P. M. (1977). *Inequality and heterogeneity: A primitive theory of social structure*, volume 7. Free Press New York.

Bo Rothstein, D. S. (2008). The state and social capital: An institutional theory of generalized trust. *Comparative Politics*, 40(4):441–459.

Bourdieu, P. (1986). *The Forms of Capital*. New York: Greenwood Press.

Brañas-Garza, P., Durán, M. A., and Espinosa, M. P. (2012). Favouring friends. *Bulletin of Economic Research*, 64(2):172–178.

Brandts, J. and Solà, C. (2010). Personal relations and their effect on behavior in an organizational setting: An experimental study. *Journal of economic behavior & organization*, 73(2):246–253.

Bratton, M. (2007). Formal versus informal institutions in africa. *Journal of Democracy*, 18(3):96–110.

Buchan, N. and Croson, R. (2004). The boundaries of trust: Own and others' actions in the us and china. *Journal of Economic Behavior & Organization*, 55(4):485–504.

Buchan, N. R., Johnson, E. J., and Croson, R. T. (2006). Let's get personal: An international examination of the influence of communication, culture and social distance on other regarding preferences. *Journal of Economic Behavior & Organization*, 60(3):373–398.

Buonanno, P., Montolio, D., and Vanin, P. (2009). Does social capital reduce crime? *Journal of Law and Economics*, 52(1):145–170.

Burt, R. S. (1982). Toward a structural theory of action.

Burt, R. S. (2000). The network structure of social capital. *Research in organizational behavior*, 22:345–423.

Calvo-Armengol, A. and Jackson, M. O. (2004). The effects of social networks on employment and inequality. *American economic review*, pages 426–454.

Camerer, C. (2010). *Behavioral game theory*. New Age International.

Cardenas, J. C. and Carpenter, J. (2008). Behavioural development economics: lessons from field labs in the developing world. *The Journal of Development Studies*, 44(3):311–338.

Çayli, B. (2010). Social networks of the italian mafia; the strong and weak parts. *CEU Political Science Journal*, (03):382–412.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American journal of sociology*, pages S95–S120.

Coleman, J. S. (1994). *Foundations of social theory*. Harvard university press.

Cox, J. C. (2004). How to identify trust and reciprocity. *Games and economic behavior*, 46(2):260–281.

Degli Antoni, G. (2009). Intrinsic vs. extrinsic motivations to volunteer and social capital formation. *Kyklos*, 62(3):359–370.

Delhey, J., Newton, K., and Welzel, C. (2011). How general is trust in "most people"? solving the radius of trust problem. *American Sociological Review*, 76(5):786–807.

D'Hernoncourt, J. and Méon, P.-G. (2012). The not so dark side of trust: Does trust increase the size of the shadow economy? *Journal of economic behavior & organization*, 81(1):97–121.

Di Falco, S. and Bulte, E. (2011). A dark side of social capital? kinship, consumption, and savings. *Journal of Development Studies*, 47(8):1128–1151.

Dinesen, P. T. (2013). Where you come from or where you live? examining the cultural and institutional explanation of generalized trust using migration as a natural experiment. *European sociological review*, 29(1):114–128.

DiPasquale, D. and Glaeser, E. L. (1999). Incentives and social capital: are homeowners better citizens? *Journal of urban Economics*, 45(2):354–384.

Duranton, G., Rodríguez-Pose, A., and Sandall, R. (2009). Family types and the persistence of regional disparities in europe. *Economic Geography*, 85(1):23–47.

Durlauf, S. N. (2002). On the empirics of social capital*. *The economic journal*, 112(483):F459–F479.

Ekman, J. and Amnå, E. (2012). Political participation and civic engagement: Towards a new typology. *Human Affairs*, 22(3):283–300.

Ermisch, J. and Gambetta, D. (2010). Do strong family ties inhibit trust? *Journal of Economic Behavior & Organization*, 75(3):365–376.

Ermisch, J. F. and Francesconi, M. (2001). Family structure and children's achievements. *Journal of Population Economics*, 14(2):249–270.

Etang, A., Fielding, D., and Knowles, S. (2011). Does trust extend beyond the village? experimental trust and social distance in cameroon. *Experimental Economics*, 14(1):15–35.

Fafchamps, M. and Lund, S. (2003). Risk-sharing networks in rural philippines. *Journal of development Economics*, 71(2):261–287.

Fehr, E., Fischbacher, U., von Rosenbladt, B., Schupp, J., and Wagner, G. G. (2003). A nation-wide laboratory: Examining trust and trustworthiness by integrating behavioral experiments into representative surveys. *IZA Discussion paper series*.

Fershtman, C. and Gneezy, U. (2001). Discrimination in a segmented society: An experimental approach. *Quarterly Journal of Economics*, pages 351–377.

Fershtman, C., Gneezy, U., and Verboven, F. (2005). Discrimination and nepotism: The efficiency of the anonymity rule. *The Journal of Legal Studies*, 34(2):pp. 371–396.

Fine, B. (2001). Social capital and realm of the intellect. *Economic and Political Weekly*, pages 741–745.

Foley, M. W. and Edwards, B. (1999). Is it time to disinvest in social capital? *Journal of public policy*, 19(02):141–173.

Freitag, M. and Traunmüller, R. (2009). Spheres of trust: An empirical analysis of the foundations of particularised and generalised trust. *European Journal of Political Research*, 48(6):782–803.

Fukuyama, F. (1995). Trust: The social virtues and the creation of prosperity. *World and I*, 10:264–268.

Fukuyama, F. (2001). Social capital, civil society and development. *Third world quarterly*, 22(1):7–20.

Galasso, V. and Profeta, P. (2013). From family culture to welfare state design. *Centre for Household, Income, Labour and Demographic Economics (CHILD)-CCA*.

Gambetta, D. (1996). *The Sicilian Mafia: the business of private protection*. Harvard University Press.

Gennetian, L. A. (2005). One or two parents? half or step siblings? the effect of family structure on young children's achievement. *Journal of Population Economics*, 18(3):415–436.

Gittell, R. and Vidal, A. (1998). *Community organizing: Building social capital as a development strategy*. Sage publications.

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

Goudge, J. and Gilson, L. (2005). How can trust be investigated? drawing lessons from past experience. *Social Science & Medicine*, 61(7):1439–1451.

Granovetter, M. S. (1973). The strength of weak ties. *American journal of sociology*, pages 1360–1380.

Greif, A. and Tabellini, G. (2010). Cultural and institutional bifurcation: China and europe compared. *The American Economic Review*, 100(2):135–140.

Greif, A. and Tabellini, G. (2015). The clan and the corporation: Sustaining cooperation in china and europe.

Grießhaber, N. and Geys, B. (2012). Civic engagement and corruption in 20 european democracies. *European Societies*, 14(1):57–81.

Guiso, L., Sapienza, P., and Zingales, L. (2008). Alfred marshall lecture social capital as good culture. *Journal of the European Economic Association*, 6(2-3):295–320.

Guiso, L., Sapienza, P., and Zingales, L. (2009). Cultural biases in economic exchange? *The Quarterly Journal of Economics*, pages 1095–1131.

Hardin, R. (2002). *Trust and trustworthiness*. Russell Sage Foundation.

Hayoz, N. and Sergeyev, V. (2003). *Social networks in Russian politics*, volume 20, page 46. London: Routledge.

Helliwell, J. F. (2007). Well-being and social capital: Does suicide pose a puzzle? *Social indicators research*, 81(3):455–496.

Hibbing, J. R. and Theiss-Morse, E. (2002). *Stealth democracy: Americans' beliefs about how government should work*. Cambridge University Press.

Hoff, K. and Sen, A. (2005). The kin system as a poverty trap? *World Bank Policy Research Working Paper*, (3575).

Holm, H. J. and Danielson, A. (2005). Tropic trust versus nordic trust: Experimental evidence from tanzania and sweden*. *The Economic Journal*, 115(503):505–532.

Hooghe, M. (2003). Participation in voluntary associations and value indicators: The effect of current and previous participation experiences. *Nonprofit and voluntary sector quarterly*, 32(1):47–69.

Iglič, H. (2010). Voluntary associations and tolerance: An ambiguous relationship. *American Behavioral Scientist*, 53(5):717–736.

Ioannides, Y. M. and Loury, L. D. (2004). Job information networks, neighborhood effects, and inequality. *Journal of economic literature*, pages 1056–1093.

Isham, J., Kolodinsky, J., and Kimberly, G. (2006). The effects of volunteering for nonprofit organizations on social capital formation: Evidence from a statewide survey. *Nonprofit and Voluntary Sector Quarterly*, 35(3):367–383.

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

Johnson, N. D. and Mislin, A. (2012). How much should we trust the world values survey trust question? *Economics Letters*, 116(2):210 – 212.

Johnson, N. D. and Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology*, 32(5):865–889.

Knack, S. and Keefer, P. (1997). Does social capital have an economic payoff? a cross-country investigation. *The Quarterly journal of economics*, pages 1251–1288.

La Ferrara, E. (2007). Family and kinship ties in development: An economist's perspective. In *5th AFD-EUDN Conference, Paris, December*.

La Ferrara, E. and Milazzo, A. (2014). Customary norms, inheritance, and human capital: evidence from a reform of the matrilineal system in ghana. *CEPR Discussion Paper No. DP10159*.

La Porta, R., Lopez-de Silanes, F., Shleifer, A., and Vishny, R. (1999). The quality of government. *Journal of Law, Economics, and organization*, 15(1):222–279.

Lerman, R. I. (1996). The impact of the changing us family structure on child poverty and income inequality. *Economica*, pages S119–S139.

Lin, N., Cook, K. S., and Burt, R. S. (2001). *Social capital: Theory and research*. Transaction Publishers.

Macy, M. W. and Skvoretz, J. (1998). The evolution of trust and cooperation between strangers: A computational model. *American Sociological Review*, pages 638–660.

Marien, S. and Hooghe, M. (2011). Does political trust matter? an empirical investigation into the relation between political trust and support for law compliance. *European Journal of Political Research*, 50(2):267–291.

McEvily, B., Radzevick, J. R., and Weber, R. A. (2012). Whom do you distrust and how much does it cost? an experiment on the measurement of trust. *Games and Economic Behavior*, 74(1):285 – 298.

McKnight, D. H., Cummings, L. L., and Chervany, N. L. (1998). Initial trust formation in new organizational relationships. *Academy of Management review*, 23(3):473–490.

Miller, A. S. and Mitamura, T. (2003). Are surveys on trust trustworthy? *Social Psychology Quarterly*, pages 62–70.

Naef, M. and Schupp, J. (2009). Measuring trust: experiments and surveys in contrast and combination. *IZA discussion papers*.

Nannicini, T., Stella, A., Tabellini, G., and Troiano, U. (2013). Social capital and political accountability. *American Economic Journal: Economic Policy*, 5(2):222–250.

Newton, K. (1997). Social capital and democracy. *American behavioral scientist*, 40(5):575–586.

Newton, K. and Zmerli, S. (2011). Three forms of trust and their association. *European Political Science Review*, 3(02):169–200.

Pearce, J. L., Branyiczki, I., and Bigley, G. A. (2000). Insufficient bureaucracy: Trust and commitment in particularistic organizations. *Organization Science*, 11(2):148–162.

Pellizzari, M. (2010). Do friends and relatives really help in getting a good job? *Industrial & Labor Relations Review*, 63(3):494–510.

Pérez-González, F. (2006). Inherited control and firm performance. *The American economic review*, pages 1559–1588.

Platteau, J.-P. (2000). *Institutions, social norms, and economic development*, volume 1. Psychology Press.

Portes, A. (2000). Social capital: Its origins and applications in modern sociology. *LESSER, Eric L. Knowledge and Social Capital. Boston: Butterworth-Heinemann*, pages 43–67.

Putnam, H. (1995). Are moral and legal values made or discovered? *Legal Theory*, 1(01):5–19.

Putnam, R. D. (2001). *Bowling alone: The collapse and revival of American community*. Simon and Schuster.

Quintelier, E., Stolle, D., and Harell, A. (2011). Politics in peer groups: Exploring the causal relationship between network diversity and political participation. *Political Research Quarterly*, page 1065912911411099.

Reeskens, T. and Hooghe, M. (2008). Cross-cultural measurement equivalence of generalized trust. evidence from the european social survey (2002 and 2004). *Social Indicators Research*, 85(3):515–532.

Robbins, B. G. (2012). Institutional quality and generalized trust: A nonrecursive causal model. *Social Indicators Research*, 107(2):235–258.

Robbins, B. G. (2016). From the general to the specific: How social trust motivates relational trust. *Social Science Research*, 55:16–30.

Rosenberg, M. (1956). Misanthropy and political ideology. *American sociological review*, pages 690–695.

Rosenfeld, R., Baumer, E. P., and Messner, S. F. (2001). Social capital and homicide. *Social Forces*, 80(1):283–310.

Rothstein, B. (2005). *Social traps and the problem of trust*. Cambridge University Press.

Rothstein, B. and Stolle, D. (2008). The state and social capital: An institutional theory of generalized trust. *Comparative politics*, pages 441–459.

Rudolph, T. J. and Evans, J. (2005). Political trust, ideology, and public support for government spending. *American Journal of Political Science*, 49(3):660–671.

Sabatini, F. (2009a). Il capitale sociale nelle regioni italiane: un'analisi comparata. *Rivista di Politica Economica*, 99(2):167–220.

Sabatini, F. (2009b). The relationship between trust and networks. an exploratory empirical analysis. *Economics Bulletin*, 29(2):661–672.

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

Sarracino, F. (2010). Social capital and subjective well-being trends: Comparing 11 western european countries. *The Journal of Socio-Economics*, 39(4):482–517.

Schechter, L. (2007). Traditional trust measurement and the risk confound: An experiment in rural paraguay. *Journal of Economic Behavior & Organization*, 62(2):272–292.

Sobel, J. (2002). Can we trust social capital? *Journal of Economic Literature*, 40(1):139–154.

Sønderskov, K. M. and Dinesen, P. T. (2015). Trusting the state, trusting each other? the effect of institutional trust on social trust. *Political Behavior*, pages 1–24.

Stolle, D. and Harell, A. (2013). Social capital and ethno-racial diversity: Learning to trust in an immigrant society. *Political Studies*, 61(1):42–66.

Stolle, D. and Lewis, J. (2002). Social capital–an emerging concept. *Key concepts in gender and European social politics*.

Tabellini, G. (2010). Culture and institutions: economic development in the regions of europe. *Journal of the European Economic Association*, 8(4):677–716.

Tan, S. J. and Tambyah, S. K. (2011). Generalized trust and trust in institutions in confucian asia. *Social indicators research*, 103(3):357–377.

Thöni, C., Tyran, J.-R., and Wengström, E. (2012). Microfoundations of social capital. *Journal of Public Economics*, 96(7):635–643.

Todd, E. and Garrioch, D. (1985). *The explanation of ideology: Family structures and social systems*. Blackwell Oxford.

Torgler, B. (2006). The importance of faith: Tax morale and religiosity. *Journal of Economic Behavior & Organization*, 61(1):81–109.

Uslaner, E. and Conley, R. S. (2003). Civic engagement and particularized trust. *American Politics Research*, 31(4):331–360.

Uslaner, E. M. (2002). *The moral foundations of trust*. Cambridge University Press.

Uslaner, E. M. (2008a). The foundations of trust: macro and micro. *Cambridge Journal of Economics*, 32(2):289–294.

Uslaner, E. M. (2008b). Trust as a moral value. *The handbook of social capital*, pages 101–121.

van Deth, J. W. (2010). Participation in voluntary associations: dark shades in a sunny world? *American Behavioral Scientist*, 53(5):640–656.

van Rijn, F., Bulte, E., and Adekunle, A. (2012). Social capital and agricultural innovation in sub-saharan africa. *Agricultural Systems*, 108:112–122.

Vivian Lei, F. V. (2010). In-group versus out-group trust: The impact of income inequality. *Southern Economic Journal*, 76(4):1049–1063.

Weber, M. (1904). *The Protestant Ethic and the Spirit of Capitalism*. New York: Scribner's Press.

Wolf, C. (1955). Institutions and economic development. *The American Economic Review*, pages 867–883.

Wollebaek, D. and Selle, P. (2002). Does participation in voluntary associations contribute to social capital? the impact of intensity, scope, and type. *Nonprofit and voluntary sector quarterly*, 31(1):32–61.

Wollebæk, D. and Strømsnes, K. (2007). Voluntary associations, trust, and civic engagement: A multilevel approach. *Nonprofit and Voluntary Sector Quarterly*.

Woolcock, M. (2001). The place of social capital in understanding social and economic outcomes. *Canadian journal of policy research*, 2(1):11–17.

Yip, W., Subramanian, S., Mitchell, A. D., Lee, D. T., Wang, J., and Kawachi, I. (2007). Does social capital enhance health and well-being? evidence from rural china. *Social science & medicine*, 64(1):35–49.

CHAPTER 2

The Effect of Particularism on Corruption: Theory and Empirical Evidence [1]

[1]This chapter is based on a joint work with Luca Stanca (DEMS, University of Milan-Bicocca). The authors wish to thank Eric Uslaner, Francesco Sarracino, participants to the Laboratory for Comparative Social Research workshop, participants to the workshop of Social Economy for young economists and participants to the 56th Annual Conference of the Italian Economic Association for useful comments and inputs. A version of this chapter has been published as: Valentina Rotondi, Luca Stanca, The effect of particularism on corruption: Theory and empirical evidence, Journal of Economic Psychology, Volume 51, December 2015, Pages 219-235

2.1 Introduction

It is widely recognized that corruption has significant adverse effects on economic development and growth (Mauro, 1995, 1997), social equality (Gupta et al., 2002) and several other relevant socio-economic phenomena (Spector, 2005). As a consequence, the search for effective policies against corruption plays a key role in the agenda of policy makers throughout the world. However, despite a growing body of literature on corruption, a full understanding of its determinants is yet to come. This paper investigates, both theoretically and empirically, the role played by the cultural norm of particularism, as opposed to universalism, as a determinant of corruption at the individual level.

A large number of studies has focused on the determinants of corruption at the macro level, showing that corruption is generally lower in countries characterized by centralized government, long lasting democracy and open-market economy (Treisman, 2000), lower ethnic heterogeneity (La Porta et al., 1999), fiscal decentralization (Fisman and Gatti, 2002), higher share of Protestants (Treisman, 2000), free press (Brunetti and Weder, 2003) and higher share of women in government (Dollar et al., 2001; Swamy et al., 2001).[2] These studies are generally based on cross-country aggregate data, and therefore do not address the role played by individual characteristics and context for the decision to engage in corruption.

With the recent availability of micro-level surveys that include specific questions on acts of corruption, such as offering or accepting bribes, a relatively smaller number of studies has turned to the determinants of corruption at the individual level (e.g., Dong et al., 2012; Dong and Torgler, 2009; Mocan, 2008; Guerrero and Rodríguez-Oreggia, 2008; Swamy et al., 2001; Torgler and Valev, 2006, 2010) . Within this micro-level literature, most studies have focused on the monetary incentives and disincentives for engaging in corruption, despite the existence of a growing literature indicating that cultural factors also play an important role for economic behavior (e.g., Alesina

[2]See Rose-Ackerman (2007) and Treisman (2007) for comprehensive reviews of this literature.

and Glaeser, 2004; Fernández and Fogli, 2009; Guiso et al., 2006). Our paper contributes to the literature on the individual-level determinants of corruption by investigating the role played by the cultural norm of particularism vs. universalism.[3]

In philosophy, the difference between particularism and universalism relates to the sphere of morality. Particularism implies that there are no general moral principles and that the moral status of an action is context-dependent (Dancy, 1993). Conversely, universalism implies that principles have to be conceived abstractly, so that they should guide individuals to do the right action in every circumstance (Kant, 1788). In political science, particularism has been defined as *"the policymakers' ability to further their career by catering to narrow interests rather than to broader national interests"* (Carey and Shugart, 1995). Political particularism stands in contrast to political universalism and its emphasis on universal rights and separation of powers.

The sociological literature extends the philosophical definition to general rules of behavior, while adding the dimension of social interactions. Parsons and Shils (1951) characterize the universalism-particularism continuum as a pattern of attitudes and behaviors typical of specific groups that, in turn, guide individual behavior. Universalism implies that correct behavior can always be defined and applied, irrespective of the context. Particularism, instead, implies that relationships come ahead of abstract social codes, so that norms and appropriate behavior are context-dependent and can occasionally be broken. To put it otherwise, while for universalists general codes of conduct take precedence over the needs and claims of friends and other personal relations, particularists tend to focus more on friendship and personal relationships than on formal rules and laws.[4]

While the attention for the effects of particularism vs. universalism is relatively recent in the economic literature, in sociology particularism has

[3]A number of studies have shown that, at the aggregate level, national culture (i.e., Davis and Ruhe, 2003; Husted and y de Estudios, 1999; Barr and Serra, 2010) , religion and religiosity (La Porta et al., 1999; Treisman, 2000), and family values (Lipset and Lenz, 2000) are important determinants of corruption.

[4]The definition by Parsons and Shils (1951) has been used by Hampden-Turner and Trompenaars (1997) to develop a model of culture defining a set of seven cultural dimensions that include the universalism-particularism continuum.

often been associated with more widespread informal institutions (Mungiu-Pippidi, 2005), lower civic mindedness and higher corruption (Lumby, 2006). This relationship, however, is theoretically underdeveloped, since the precise mechanisms underlying the causal link between particularism and corruption have not been fully understood. In order to fill this gap, this paper develops and tests empirically a model of collusive bribery, focusing on the role played by the cultural norm of particularism.[5]

In our framework, the act of offering or demanding a bribe is perceived as inappropriate, according to a commonly held social norm, thus resulting in a psychological cost for the agents involved. We assume that this psychological cost is lower for particularistic individuals, defined as in Parsons and Shils (1951), since they are less sensitive to the burden imposed by corruption on society, as in Uslaner (2002). In this perspective, social relations in particularistic societies rely on strong ties informed by principles of tradition, conformity and benevolence inside small circles of related people (i.e., members of the family, friends, members of the clan). Outside this small network, selfish behavior is considered morally acceptable (Tabellini, 2010). Therefore, as shown in Baron (2010), particularists can overcome the incentive to free-ride on others or to impose negative externalities to other people when they interact with people who are close to them.[6]

We test the predictions of the model by using individual-level data for 25 countries from the European Social Survey. The findings indicate that, controlling for a wide set of individual characteristics, particularism increases the probability of offering a bribe. This result is robust to alternative definitions of particularism, specifications of the model and econometric techniques that take into account the potential endogeneity of cultural norms. Overall, our findings indicate that there is a causal link between particularism and corruption at the individual level. As a consequence, policies aimed at favoring

[5]Collusive bribery is defined as an illegal transaction that is beneficial to both the briber and the bribee, and is therefore particularly difficult to deter (Ryvkin and Serra, 2012).

[6]This assumption is corroborated by research suggesting that closed and socially exclusive networks are a fertile ground for corruption (e.g., Lambsdorff, 2008; Lambsdorff et al., 2004; Tonoyan, 2003) and that informal institutions are more widespread in particularistic rather than universalistic societies.

universalism can provide an effective tool to reduce corruption.

The rest of the paper is organized as follows. Section 2.2 reviews the related literature. Section 2.3 presents the theoretical framework. Section 2.4 describes the data and methods. Section 2.5 presents the results. Section 2.6 concludes.

2.2 Related Literature

Our work contributes to the relatively small body of empirical literature on the determinants of corruption at the individual level. Using micro-level data for 49 countries, Mocan (2008) finds that high-income individuals, males and people living in larger cities, are more likely to be asked for a bribe. At the country-level, the paper shows that the extent of corruption, as reported by citizens, is strongly correlated with indices of corruption perception. Guerrero and Rodríguez-Oreggia (2008) study the socio-demographic and institutional factors affecting corruption in Mexico, showing that weak enforcement of law, higher education and higher income levels are associated with a higher propensity to bribe. Tavits (2010) studies the relationship between the act of offering a bribe and its justifiability. Using original survey data for Estonia, she finds that both public officials and citizens are more likely to engage in corruption when they do not view corruption as wrong, and when they perceive that corrupt behavior is widespread among their peers. Lee and Guven (2013) use micro-level data from the European Social Survey to focus on masculinity, gender roles and risk preferences as cultural traits affecting the individual propensity to engage in corruption. Their results suggest that risk-seeking individuals are significantly more likely to offer and to be asked for a bribe, and that promoting gender equality might lead to less corruption.[7]

Given the hidden nature of corruption transactions, empirical studies

[7]Svensson (2003), using data on bribe payments of Ugandan firms, finds that the incidence of corruption can be explained by the variation in policies and regulations across industries. By combining data on corruption with financial information from the surveyed firms, he shows that firms' ability to pay and refusal power can explain a large part of the variation in bribes across graft-reporting firms.

at the micro-level are particularly difficult to implement (Lee and Guven, 2013). Recently, scholars have therefore turned to experimental methods to study the motivations driving individuals to engage in corruption (see Serra and Wantchekon (2012) for a comprehensive review). Some of these experimental studies have focused on culture to explain corruption (e.g., Cameron et al., 2009; Alatas et al., 2009; Barr and Serra, 2010) . The findings in this literature generally indicate that culture can only partially explain corruption behavior at the individual level.

Another strand of related literature focuses on the micro-level determinants of justifiability of corruption. Swamy et al. (2001) study the relationship between gender and corruption. Based on World Values Surveys data for over 90,000 individuals in 49 countries, they find that women are significantly less likely to condone corruption.[8] Torgler and Valev (2006) use World Values Survey and European Values Survey data to study the effect of age on acceptability of corruption, finding a positive and significant effect of age, but no cohort effect. Dong and Torgler (2009) find that a higher level of political interest is associated with a lower acceptability of corruption and with a lower level of perceived corruption. Torgler and Valev (2010) investigate whether attitudes towards corruption and tax evasion vary systematically with gender and, more specifically, whether gender differences decrease as men and women face similar opportunities for illicit behavior. Their results indicate that aversion to corruption and tax evasion is significantly stronger for women.

A third strand of related literature investigates the relationship between social capital and corruption. The mechanism through which social capital may affect corruption is twofold. On the one hand, higher levels of social capital generally imply a stronger moral obligation for public officials to abstain from acting opportunistically. On the other hand, social capital may favor corruption, since belonging to a closed network reduces informational asymmetries, thus favoring collusive behavior. At the empirical level, high levels of interpersonal trust have been found to reduce corruption (e.g., Bjørnskov,

[8]Using survey data for 350 Georgian firms, they also find that firms owned or managed by women are significantly less likely to be involved in bribing.

2006; La Porta et al., 1999; Uslaner, 2013), whereas high levels of bonding social capital (Putnam, 2001) positively affect corruption, by leading those within a privileged network to feel that they can act illegally with impunity (Lipset and Lenz, 2000). Harris (2007) shows that indicators of strong ties, family orientations and particularized trust are significantly associated to higher levels of corruption. This suggests that the adverse effect on corruption is specific to generalized trust, as opposed to particularized trust (e.g., Warren, 2004; Uslaner, 2013).

The distinction between particularism and universalism thus plays a key role for understanding the determinants of corruption. While the sociological literature on particularism and corruption is well developed (e.g., Lambsdorff et al., 2004; Lambsdorff, 2008; Tonoyan, 2003), within economics there are relatively few papers investigating this relationship, mainly confined to the experimental setting (e.g., Fong and Luttmer, 2011; Eckel and Grossman, 2005). In a recent paper, De Blasio et al. (2014) show that three measures of particularism (in trusting behavior, political participation and associational activity) are positively related to each other and negatively related to several indicators of universalism. To the best of our knowledge, our work is the first to provide an economic analysis of the effects of particularism on corruption.

2.3 Theory

In our theoretical framework, corruption is viewed as the outcome of a social interaction, rather than contract design or information sharing (Chang and Lai, 2002). We consider a society in which risk neutral citizens and public officials meet randomly[9] and interact for the provision of a good or a service. All agents have the option to engage in bribery with their transaction partner, provided that both parties accept the illegal transaction. Each agent can be either particularist or universalist.[10] Each agent knows his own type, but

[9] As in Bowles and Garoupa (1997), Chang and Lai (2002) and Ryvkin and Serra (2012), our model focuses on a setting of "casual corruption", i.e. situations in which there is no repeated interaction between the briber and the bribee.

[10] We make the assumption that the particularism vs. universalism continuum can be treated as a binary outcome for the sake of simplicity. The results presented below are

does not know the type of his counterpart. Citizens are particularist with probability γ and universalist with probability $1-\gamma$, while public officials are particularist with probability π and universalist with probability $1-\pi$. Since all transactions involve a citizen and a public official, there are equal numbers of citizens and public officials in the economy and their populations are normalized to unity.[11]

A social norm against corruption exists among both citizens and public officials, so that violating this norm implies a psychological cost. As in Akerlof (1980), this disutility depends not only on the individual's sensitivity to the norm, but also on his perceived level of widespread corruption. In particular, when corruption is low, the cost of violating the social norm is higher, so that individuals tend to engage less in corruption. On the other hand, when corruption is widespread, the cost of violating the social norm is lower and more individuals engage in corruption (Torgler, 2003). We assume that the disutility of corruption also depends on the perceived cost imposed by corruption on society. Particularists, characterized by *limited morality* as in Tabellini (2010), bear a lower cost, relative to universalists, when engaging in corruption. As a consequence, they are less reluctant to impose a burden on the society or to free-ride on others.

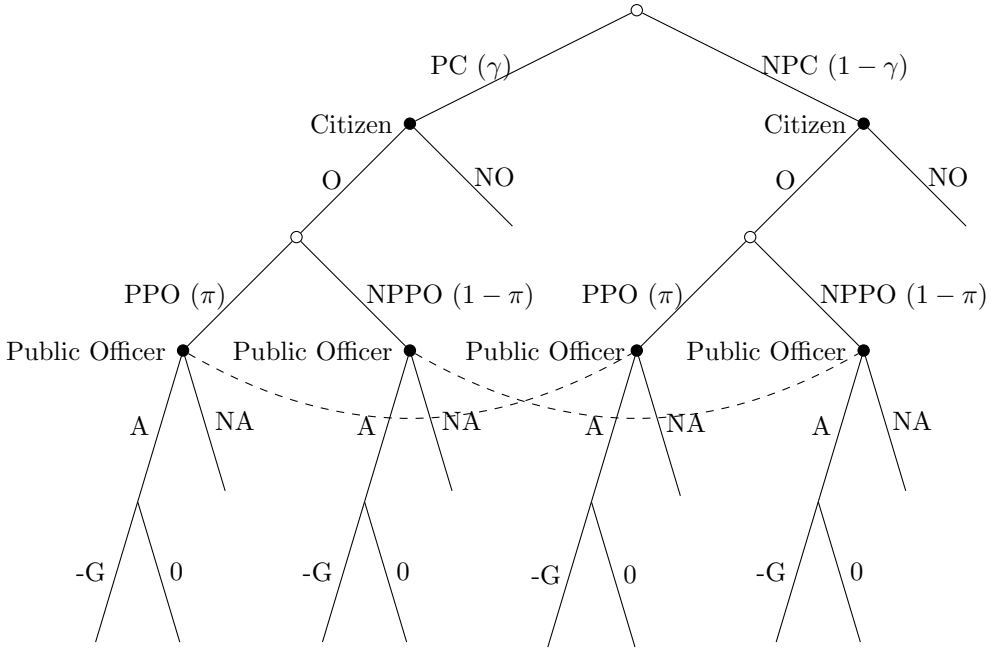
In order to prevent illegal transactions, the authority conducts random audits, so that agents engaging in corruption are jointly punished when discovered. The two agents engaging in corruption share the same probability of detection (q) and penalty (G). The timing of the corruption problem is described in Figure 2.1.

Let us denote the endowments for the citizen and the public official as Y_c and Y_p, respectively, the bribe as B, and the gain from corruption for the citizen as K. In the first stage, the citizen internalizes the possibility of collusion with the public official and evaluates the benefits and costs of

qualitatively unchanged under a more general definition of particularism.

[11]The assumption of equal number of citizens and public officials is made for simplicity, since it allows us to interpret the shares of those offering or accepting bribes (α and β, respectively) as individual-level probabilities. The same assumption is made in Boadway et al. (2002) and, implicitly, in Ryvkin and Serra (2012). It is also equivalent to assume, as in Chang and Lai (2004), that a seller faces N customers with N normalized to unity.

Figure 2.1: Corruption game tree



Note: (N)PC=(Non-)Particularist citizen, (N)PPO=(Non-)Particularist Public Officer, (N)O=(Does not) Offer bribe, (N)A=(Does not) Accept bribe. γ = probability that Citizen is particularist, π = probability that Public Officer is particularist, G = penalty if detected (with probability q).

corruption. If he chooses not to offer a bribe, then his payoff is Y_c, whereas if he chooses to offer a bribe his payoff depends on his own type, the choice made by the public official, and the probability of being discovered. If the public official accepts the bribe, the citizen's expected payoff is

$$\begin{aligned} &q(Y_c + K - C_c - B - G) + (1-q)(Y_c + K - C_c - B) = \\ &Y_c + K - C_c - B - qG \end{aligned} \tag{2.1}$$

The psychological cost of corruption is defined as

$$C_c = C_c(\mu, \theta, C) = \frac{\mu}{\theta} C \tag{2.2}$$

where μ is the subjective sensitivity to the social norm. The density function of μ, denoted by $f(\mu)$, is assumed to be uniform with support on $[0, 1]$. The

parameter θ, with $0 < \theta \leq 1$, denotes the perception of overall corruption. C is the perceived cost imposed by corruption on society. It is equal to L if the citizen is particularist (with probability γ) and to H if the citizen is universalist (with probability $1 - \gamma$), with $L < H$.

If a citizen decides to offer a bribe and the public official does not accept, the citizen's payoff is Y_c. Given that the matching between a citizen and a public official is isolated and random, the probability of successful collusion faced by a citizen is given by the fraction of collaborative public officials, denoted with β. By internalizing this success rate in his decision, the citizen's expected payoff is:

$$\Lambda = (1 - \beta)(Y_c) + \beta(Y_c + K - C_c - B - qG) \tag{2.3}$$

In the second stage, the public official decides whether or not to accept the citizen's proposal for collusion. If he refuses, then his payoff is Y_p, whereas if he accepts his payoff is

$$\begin{aligned} &q(Y_p + B - G - C_p) + (1 - q)(Y_p + B - C_p) = \\ &Y_p + B - C_p - qG \end{aligned} \tag{2.4}$$

Thus, if the public official accepts, he will obtain the bribe B, but will also suffer a psychological cost C_p,[12] arising from the violation of the social norm, where

$$C_p = C_p(\varepsilon, \theta, C) = \frac{\varepsilon}{\theta} C \tag{2.5}$$

As for the citizen's cost above, C_p is assumed to depend on the individual-specific sensitivity to the social norm (ε), uniformly distributed on $[0, 1]$, the perception of overall corruption (θ), and the perceived cost imposed by corruption on society (C), that is L for a particularist (with probability π) and H for a universalist (with probability $1 - \pi$).

In the third stage, the bribe amount B is determined by bargaining be-

[12]The moral cost of the agent who takes the initiative to break the social norm is likely to be higher than that of the counterpart. However, it is possible to show that introducing such heterogeneity in the psychological cost of engaging in bribery does not alter the key predictions of the model.

tween the public official and the citizen.[13] The model can be solved backwards, starting from the bargaining game in the third stage.

2.3.1 Bargaining Over the Bribe Amount

In the last stage of the game, the size of the bribe B is determined by Nash bargaining between the public official and the citizen. The amount of the bribe in equilibrium is therefore determined by maximizing the product of each player's gain from making a deal, with the bargain being feasible if and only if the net gain from bargaining is positive for each party:

$$\begin{aligned} &\max_{B} \left[K - C_c - B - qG\right]\left[B - C_p - qG\right] \\ &\quad \text{subject to} \\ &\quad K \geq B + qG + C_c \\ &\quad B \geq qG + C_p \end{aligned} \tag{2.6}$$

The feasible bribery set is therefore

$$qG + C_p \leq B \leq K - qG - C_c \tag{2.7}$$

and the equilibrium bribe with equal bargaining power is:

$$B^* = B^*(K, C_p, C_c) = \frac{1}{2}\left[K - C_c + C_p\right] \tag{2.8}$$

i.e.

$$B^* = B^*(K, H, L, \gamma, \pi, \varepsilon, \mu) = \frac{1}{2}\left[K - \frac{\mu}{\theta}((1-\gamma)H + \gamma L) + \frac{\varepsilon}{\theta}((1-\pi)H + \pi L)\right]. \tag{2.9}$$

From the expression above, $\frac{\partial B^*}{\partial C_p} > 0$ indicates that the higher the cost for the public official of taking part in a corrupt transaction (C_p), the larger the equilibrium bribe. $\frac{\partial B^*}{\partial K} > 0$ shows that the higher the gain from corruption (K), the higher the size of the bribe. Finally, the higher the psychological

[13] Bribery is often the result of bargaining between public officials and clients (either firms or citizens). This explains the within-country and within-sector variation in both the frequency of corrupt transactions and the size of the bribes paid (Svensson, 2003).

cost born by the corrupt citizen, the lower the size of the bribe ($\frac{\partial B^*}{\partial C_c} < 0$). Therefore, if the probability that the public official is particularist is high, the equilibrium bribe will be smaller ($\frac{\partial B^*}{\partial \pi} < 0$). On the other hand, if the probability that the citizen is particularist is high, the equilibrium bribe will be larger ($\frac{\partial B^*}{\partial \gamma} > 0$).

2.3.2 The Public Official's Decision

The public official is willing to collaborate with the corrupt citizen as long as:

$$\pi(B - C_p - qG) + (1 - \pi)(B - C_p - qG) \geq 0 \tag{2.10}$$

This expression indicates that the public official is corruptible if the net gain from corruption is positive, and therefore, given $C_p(\varepsilon, \theta, C)$, public officials with a lower value of ε are more likely to accept bribes. The critical ε^* which makes a public official indifferent between engaging in or abstaining from corruption is therefore:

$$\varepsilon^* = \theta \left[\frac{K - 2qG - C_c}{(1 - \pi)H + \pi L} \right] \tag{2.11}$$

Then, the fraction of public officials accepting bribes is equal to ε^*:

$$\beta = \int_0^{\varepsilon^*} f(\varepsilon)d\varepsilon = \varepsilon^* = \theta \left[\frac{K - 2qG - C_c}{(1 - \pi)H + \pi L} \right] = \theta \left[\frac{K - 2qG - \frac{\mu}{\theta}((1 - \gamma)H + \gamma L)}{(1 - \pi)H + \pi L} \right]$$

where $\frac{\partial \beta}{\partial q} < 0$ and $\frac{\partial \beta}{\partial G} < 0$ indicate that the higher the probability of being detected, or the higher the size of the fine, the lower the willingness to accept bribes. $\frac{\partial \beta}{\partial \pi} > 0$ suggests that the higher the probability that the public official is particularist, the higher the probability that he will accept bribes. Finally, $\frac{\partial \beta}{\partial \theta} > 0$ indicates that the more widespread corruption, the higher the probability of accepting a bribe.

2.3.3 The Citizen's Decision

Given the public official's incentives, citizens internalize the probability β that a bribe is accepted and choose whether to offer a bribe or not. Since matchings are randomly determined, and the sensitivity of the public official to the social norm cannot be observed *ex ante*, citizens do not know the exact size of the bribe before the bargain takes place. However, citizens recognize that collusion can occur if and only if the public official's sensitivity to the social norm is sufficiently small ($0 < \varepsilon < \varepsilon^*$). Thus, based on equations (2.9) and (2.2), the expected bribery amount $E(B)$, under the condition $0 < \varepsilon < \varepsilon^*$, is

$$E(B) = \frac{1}{2}\left[K + C_c + \frac{1}{\theta}E(\varepsilon|0 < \varepsilon < \varepsilon^*)((1-\pi)H + \pi L)\right] \tag{2.12}$$

with $E(\varepsilon|0 < \varepsilon < \varepsilon^*) = \frac{\varepsilon^*}{2}$ given the assumption that ε has a uniform distribution. Using equations (2.9) and (2.11), the expected bribery amount is equal to

$$E(B) = \frac{3}{4}K - \frac{3}{4}C_c - \frac{1}{2}qG. \tag{2.13}$$

A citizen will offer a bribe as long as

$$\gamma\left[(1-\beta)(Y_c) + \beta(Y_c + K - L\frac{\mu}{\theta} - B - qG) - Y_c\right] +$$

$$(1-\gamma)\left[(1-\beta)(Y_c) + \beta(Y_c + K - H\frac{\mu}{\theta} - B - qG) - Y_c\right] > 0 \tag{2.14}$$

i.e.

$$K - E(B) - qG - \frac{\mu}{\theta}\left[(1-\gamma)H + \gamma L\right] > 0 \tag{2.15}$$

Using (2.13) in (2.15), given equation (2.2), we can find the value μ^* that makes the citizen indifferent between engaging in bribery or not:

$$\mu^* = \theta\left[\frac{K - 2qG}{(1-\gamma)H + \gamma L}\right] \tag{2.16}$$

As above, given the uniform distribution of μ, we obtain:

$$\alpha = \int_0^{\mu^*} f(\mu)d\mu = \mu^* = \theta\left[\frac{K - 2qG}{(1-\gamma)H + \gamma L}\right] \tag{2.17}$$

where $\frac{\partial\alpha}{\partial\gamma} > 0$, $\frac{\partial\alpha}{\partial q} < 0$, $\frac{\partial\alpha}{\partial G} < 0$ and $\frac{\partial\alpha}{\partial\theta} > 0$. These partial effects indicate that the higher the probability of being detected, or the higher the size of the fine, the lower the probability that the citizen offers a bribe; the higher the probability that the citizen is particularist, or the higher the perception of overall corruption, the higher the probability that the citizen offers a bribe.

The model also provides explicit predictions about how the effect of particularism on offering a bribe is affected by aggregate variables, such as the overall diffusion of corruption, the level of deterrence and the pervasiveness of particularism. Taking the partial derivative of (2.17) with respect to γ, we obtain

$$\frac{\partial\alpha}{\partial\gamma} = \theta(H - L)\frac{K - 2qG}{[(1-\gamma)H + \gamma L]^2} \tag{2.18}$$

The expression in (2.18) indicates that the effect of individual particularism on the probability to offer a bribe is higher when corruption is more widespread ($\frac{\partial^2\alpha}{\partial\gamma\partial\theta} > 0$). When corruption is widespread, the social stigma associated with this behavior is smaller, so that the cost of offering a bribe is lower. Since the effects of individual particularism and overall corruption are complements, being particularist has a stronger effect on bribe offering in countries where corruption is more widespread. From (2.18) it can also be shown that $\frac{\partial^2\alpha}{\partial\gamma\partial G} < 0$ and $\frac{\partial^2\alpha}{\partial\gamma\partial q} < 0$, indicating that the effect of particularism on bribe offering is smaller in countries where deterrence is stronger.

Furthermore, $\frac{\partial^2\alpha}{\partial\gamma^2} = -2(L - H)(H - L)\theta\frac{K-2qG}{[(1-\gamma)H+\gamma L]^3} > 0$, indicating that the effect of particularism on the decision to offer a bribe is higher when the fraction of particularists in the society is higher.[14]

[14]It is possible to consider cases where it is the public official who asks the citizen for a bribe, rather than the citizen offering a bribe to the public official. A modified version of the model that considers such transactions is presented in Appendix A.

2.4 Data and Methods

Our empirical analysis is based on the second round of the European Social Survey (ESS) [15], covering 49,066 individuals in 25 countries for the period 2004-2006.[16] The questionnaire in the ESS included, among others, questions on family status, employment, well-being, health and economic morality. Three main questions on bribery were included in the questionnaire. The first is whether respondents have offered a bribe ("How often, if ever, have you offered a favor or bribe to a public official in return for their services in the last five years?"). The second is whether respondents have been asked for a bribe ("How often, if ever, a public official asked you for a favor or a bribe in return for a service in the last five years?"). The responses are coded as never = 1, once = 2, twice = 3, three or four times = 4 and five times or more = 5. For ease of interpretation, we re-coded these variables into binary outcomes with "ever offered/been asked for bribes" equal to 1 and "never" equal to 0. The third question is about bribe justification ("How wrong is that a public official asks someone for a favor or bribe in return for their services?"), with responses coded as: seriously wrong = 1, wrong = 2, a bit wrong = 3, and not wrong at all = 4. For the 25 countries in the sample, 1.7 per cent of the individuals in the sample report to have offered a bribe and about 4 per cent to have been asked for a bribe. As shown in Table 2.9 in Appendix B, there is substantial variability across countries, with Finland displaying the lowest fraction of people declaring to have offered or been asked for a bribe (0.15 per cent and 0.90, respectively) and Slovakia the highest (7.8 per cent and 14.3 per cent, respectively).

We operationalize Parsons and Shils' (1951) definition of particularism as the difference between the answer to the question "It is occasionally alright to ignore the law and do what you want" and the question "You should always strictly obey the law, even if it means missing good opportunities."[17]

[15] Edition 3.2 released on February 2, 2011

[16] Although there are 26 countries in the ESS, observations for Ukraine were excluded due to data quality problems.

[17] Both variables are coded as "Disagree strongly"=1, "Disagree"=2, "Neither disagree nor agree"=3, "Agree"=4, "Agree strongly"=5.

We define an individual as particularist if the difference between these two variables is positive (*particularism*). Therefore, any individual in the sample who reports a higher score in the first question than in the second is identified as a particularist, while all the others are considered non-particularists.

In our empirical specifications, we include among control variables the respondent's self-reported political interest, trust, religion, education measured as completed years of schooling, family income (Dong and Torgler, 2009; Torgler and Valev, 2010; Guerrero and Rodríguez-Oreggia, 2008). We also include risk tolerance and a gender cultural indicator as defined in Lee and Guven (2013). The estimated specifications include a variable accounting for individual honesty ("How much do you agree or disagree with this statement: If you want to make money, you cannot always act honestly")[18] and a variable accounting for the importance of being admired ("Now I will briefly describe some people. Please listen to each description and tell me how much each person is or is not like you. It's important to her/him to show her/his abilities. She/he wants people to admire what she/he does.") in order to control for the effects of social desirability. Moreover, in order to take into account the networks of relations that can favor collusive bribery, we included among the regressors a variable measuring the number of friends who are willing to support the respondent in case of illegal activities.[19]

In two specifications of our model, we also accounted for the perception of overall corruption by including an aggregate measure of the level of overall corruption, the Corruption Perception Index (International), or a variable measuring the fraction of individuals that in each country declared to have actually offered a bribe (Corruption Diffusion). Furthermore, in order to account for the government's ability to deter corruption and for the perception of the extent to which public power is exercised for private gain, we also added an aggregate measure of the Worldwide Governance Indicators of the

[18] The answers are coded as agree strongly=1, agree=2, neither agree nor disagree=3, disagree=4, disagree strongly=5. In the sample 28% of the respondents declared that they agree strongly, 18% that they agree, 33% that they neither agree not disagree, 14% that they disagree and the remaining 8% that they disagree.

[19] "Suppose you planned to get benefits or services you were not entitled to. How many of your friends or relatives do you think you could ask for support?."

World Bank (Control of Corruption). Table 2.1 presents descriptive statistics for all the variables used in the econometric analysis.

In order to assess the robustness of our definition of particularism, we also considered a number of alternative indicators. *Particularism2* and *particularism3* are based on an alternative threshold (the difference between the two underlying components being greater than 1 and -1, respectively). *Particularism4* is a discrete variable defined as the simple difference between the two underlying components described above. We also constructed additional indicators based on the fact that, as suggested by Uslaner (2013), particularists rely on strong family ties and are characterized by low levels of generalized trust. We therefore created a dummy variable (*particularism5*) equal to one when both *particularism* and the dummy variable indicating strong family ties are equal to one.[20] In order to account for the fact that particularists are characterized by a lower degree of generalized trust, we created a dummy variable equal to one when particularism is equal to one and respondent's generalized trust is lower than the median of the weighted sample (*particularism6*). We also compared our indicators of particularism with measures of general trust and respect for rules.[21] All the indicators of particularism are positively and significantly correlated pair-wise among them and negatively and significantly correlated with the indicators of general trust and respect for rules. This indicates that our measure of particularism is qualitatively robust to alternative definitions. In the following analysis, we thus focus on the indicator *particularism* as the key explanatory variable, while the alternative indicators of particularism are used for assessing the robustness of the results.

The theoretical model presented in Section 2.3 provides a number of relevant testable predictions. First, at the individual level, particularism decreaseses the psychological cost of bribing. As a consequence, being particularist

[20]Overall, an individual has strong family ties when he declares that: (1) A person's family ought to be his or her main priority in life and that (2) when choosing a job is important that it allows to combine work and family.

[21]As for trust, we considered the standard question "Generally speaking, would you say that most people can be trusted?", while respect for general laws was proxied by the answer to the question: "How much do you agree or disagree with this statement about how people see rules and laws: It is important to follow rules ".

Table 2.1: Descriptive statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Offered bribe	0.02	0.13	0	1	45503
Been asked for bribe	0.05	0.21	0	1	43074
Bribe wrong	0.96	0.19	0	1	45419
Particularism	0.14	0.35	0	1	44811
Particularism2	0.14	0.35	0	1	44811
Particularism3	0.51	0.5	0	1	44811
Particularism4	-1.18	1.63	-4	4	44811
Particularism5	0.05	0.22	0	1	46955
Particularism6	0.02	0.15	0	1	46955
Gender	0.46	0.5	0	1	46955
Age	46.48	18.4	13	102	46714
Age squared	2498.82	1812.09	169	10404	46714
Education	11.5	4.07	0	44	46398
Employed	0.5	0.5	0	1	46955
Married	0.53	0.5	0	1	44941
Family members	2.87	1.5	1	18	46926
Income	2.02	0.86	1	4	44418
Minority ethnic group	1.96	0.2	1	2	46015
Immigrant	0.06	0.24	0	1	46585
Son of immigrants	0.02	0.15	0	1	46585
Big city	0.2	0.4	0	1	46846
Suburbs big city	0.12	0.33	0	1	46846
Village	0.3	0.46	0	1	46846
Friends support	1.6	0.76	1	4	35532
Political interest	2.67	0.9	1	4	46803
Follow rules	3.94	1.35	1	6	43676
Trust	5.92	2.48	1	11	46762
Trust in legal system	6.18	2.62	1	11	45567
Trust in public officials	3.21	0.92	1	5	44038
Religiosity	5.89	2.99	1	11	46624
Risk	3.97	1.45	1	6	43970
Male role index	3.41	0.77	1	5	44738
Admired	3.23	1.38	1	6	43876
Honesty	3.17	1.2	1	5	45053
Corruption Diffusion	35.11	49.7	1	203	46955
COC	1.4	0.85	-0.33	2.59	46955
CPI	6.73	2.03	3.2	9.70	46955
European unification	6.28	2.67	1	11	42451
Important care nature	4.86	1.03	1	6	44051

Source: European Social Survey, 2004-2006. See Section 2.4 for a description of the variables.

increases the probability of offering a bribe and, to the extent that particularism is observable, also the probability of being asked for a bribe. Second, at the aggregate level, corruption deterrence decreases the probability of offering and being asked for a bribe, while the pervasiveness of corruption increases the probability of offering and being asked for a bribe. Third, focusing on moderating factors, the effect of individual particularism on the probability to offer a bribe is higher in countries where corruption or particularism are more widespread, while it is lower in countries where corruption deterrence is stronger.

2.5 Results

This section presents the results of the empirical analysis. We start by using univariate probit models to estimate the effect of particularism on the probability of offering and being asked for a bribe, respectively. We focus on the effect of particularism, while controlling for a wide set of control variables, at both individual and country level. Standard errors are clustered by country, in order to take into account the nested structure of the data (individual level within country level), and the use of aggregate variables in a micro-level specification. We then present the results of Instrumental Variables and Propensity Score estimation to validate the causal interpretation of our findings. Finally, we turn to a structural equation model to explicitly address the simultaneous relationship between particularism, the psychological cost of bribing and the decision to offer a bribe.

2.5.1 The Determinants of Bribing

Table 2.2 reports probit estimation results (marginal effects multiplied by 100) for the determinants of the decision to offer a bribe. We consider five different specifications, with progressively larger sets of control variables.[22]

[22]It should be observed that the use of larger sets of control variables reduces the number of observations. However, it allows us to obtain a more complete characterization of the determinants of the decision to engage in corruption.

The first specification includes only standard socio-demographic characteristics, such as age, years of schooling, gender, marital and employment status, household's composition, income and minority status. Model (2) also accounts for immigration status and living context, as well as network characteristics and political interest. Model (3) also includes variables related to trust (both general and institutional) and religiosity. Models (4) and (5) include alternative country-level indicators of the pervasiveness of corruption. Except for these last two models, all specifications also include country dummies.

Consistent with the theoretical predictions, particularism has a positive and significant effect on the probability of offering a bribe in all specifications. The size of the effect is small in absolute terms, ranging between 1.42 and 0.75 probability points, similarly to related studies based on micro data (e.g., Mocan, 2008; Lee and Guven, 2013). This reflects the fact that having offered a bribe has a very small frequency in the sample (about 1.6 per cent).[23] In relative terms, however, the size of the effect of particularism is sizeable, being larger than that of gender, employment status and all other individual characteristics. As shown in Table 2.10 in Appendix B, the key result of a positive effect of particularism on the decision to offer a bribe is robust to the use of alternative definitions of particularism, as described in Section 2.4.

Focusing on aggregate variables, higher levels of perceived corruption (i.e., higher Corruption Diffusion) are associated to a significantly higher probability of offering a bribe. Stronger corruption deterrence (i.e., higher Control of Corruption) is negatively related to the probability of offering a bribe, although not significantly. Turning to individual characteristics, in line with the literature, institutional trust (in the legal system and in public officials) is associated with a significantly lower probability of offering a bribe. Being an immigrant or a son of an immigrant is not significantly related to the probability of offering a bribe. As expected, being more honest is associated with a significantly lower probability of offering a bribe, while having a large

[23]The large size of our sample allows us to rule out the possible bias arising from the fact that our binary dependent variable describes a rare event (King and Zeng, 2001). Indeed, using a penalized maximum likelihood regression approach (Firth, 1993), the results are virtually unchanged.

number of friends supporting illegal activities is associated with a significantly higher probability of offering a bribe (Rose-Ackerman, 2001). As for gender, we find that, consistent with the literature, males are significantly more likely to offer a bribe than females (Lee and Guven, 2013).

Table 2.2: Determinants of offering a bribe, probit estimates

	(1)	(2)	(3)	(4)	(5)
Particularism (d)	1.42***	1.08***	0.75***	0.76***	0.88***
Gender (d)	0.38***	0.27**	0.18	0.22**	0.19
Age	0.01	0.01	-0.00	-0.01	-0.01
Age squared	-0.00	-0.00	0.00	0.00	0.00
Education	-0.00	0.01	0.01	0.01	0.04**
Employed (d)	-0.26***	-0.31***	-0.35***	-0.39***	-0.39***
Married (d)	-0.03	-0.05	-0.03	0.01	-0.05
Family members	-0.02	-0.03	-0.03	0.03	-0.04
Income	0.01	0.03	0.00	0.08	0.04
Minority ethnic group	-0.48***	-0.52**	-0.42*	-0.45*	-0.36
Immigrant (d)		0.25	0.30	0.13	0.39
Son of immigrants (d)		0.39	0.17	-0.00	0.13
Big city (d)		0.18	0.22*	0.39**	0.04
Suburbs big city (d)		0.20	0.21	0.18	0.09
Village (d)		0.15	0.18	0.25	0.23
Friends support		0.48***	0.39***	0.49***	0.50***
Political interest		-0.09	-0.09	-0.12	-0.07
Follow rules		-0.06*	-0.04*	-0.02	-0.00
Trust			0.01	-0.00	0.03
Trust in legal system			-0.03	-0.07**	-0.06**
Trust in public officials			-0.24***	-0.24***	-0.25***
Religiosity			0.02*	0.06***	0.00
Risk			-0.06*	-0.06	-0.06
Male role index			-0.09*	-0.12**	-0.07
Admired			-0.01	-0.03	0.04
Honesty			-0.22***	-0.24***	-0.16***
Corruption Diffusion				0.01***	
COC					-0.29
CPI					0.31
Observations	40456	29202	26645	27066	27066
Country dummies	Yes	Yes	Yes	No	No

Note: probit estimates (marginal effects multiplied by 100). Dependent variable: binary variable for having offered a bribe. (d) indicates discrete change of dummy variable from 0 to 1. Standard errors clustered by country. * denotes significance at 0.10 level (** at 0.05, *** at 0.01).

Table 2.3 reports probit estimates (marginal effects multiplied by 100) for the determinants of being asked for a bribe, using the same set of specifica-

tions and control variables as for the probability of offering a bribe. Overall, we find a positive and significant relation between particularism and being asked for a bribe. This can be interpreted as an indication that particularism, rather than being private information of the bribee, is to some extent observable (see Appendix A).

Table 2.3: Determinants of being asked for a bribe, probit estimates

	(1)	(2)	(3)	(4)	(5)
Particularism (d)	2.52***	1.86***	0.79***	0.57*	0.89***
Gender (d)	1.32***	1.14***	0.66***	0.76***	0.67***
Age	0.09**	0.08**	0.05	0.03	0.05
Age squared	-0.00***	-0.00***	-0.00*	-0.00	-0.00
Education	0.14***	0.11***	0.12***	0.10*	0.14***
Employed (d)	-0.47**	-0.41*	-0.38*	-0.41	-0.50**
Married (d)	0.17	0.45**	0.52**	0.80***	0.54**
Family members	-0.11	-0.10	-0.12	0.01	-0.14
Income	0.63***	0.82***	0.53***	0.85***	0.60***
Minority ethnic group	-1.22***	-0.84**	-0.47	-0.44	-0.43
Immigrant (d)		0.60	0.76**	0.56	1.43**
Son of immigrants (d)		-0.10	-0.13	-0.33	0.16
Big city (d)		0.82***	0.83***	1.95***	0.83**
Suburbs big city (d)		0.26	0.26	0.47	0.23
Village (d)		-0.46	-0.32	-0.25	-0.27
Friends support		1.08***	0.81***	0.88***	0.92***
Political interest		-0.60***	-0.54***	-0.65***	-0.55***
Follow rules		-0.10	-0.06	-0.00	-0.01
Trust			-0.07	-0.17***	-0.05
Trust in legal system			-0.07*	-0.09	-0.06
Trust in public officials			-1.37***	-1.48***	-1.44***
Religiosity			0.05*	0.19***	0.06*
Risk			-0.23***	-0.32***	-0.30***
Male role index			-0.06	-0.21	0.01
Admired			-0.17**	-0.32***	-0.11
Honesty			-0.46***	-0.63***	-0.45***
Corruption Diffusion				0.03***	
COC					-0.40
CPI					1.21**
Observations	38451	28444	26281	26281	26281
Country dummies	Yes	Yes	Yes	No	No

Note: probit estimates (marginal effects multiplied by 100). Dependent variable: binary variable for having being asked for a bribe. (d) indicates discrete change of dummy variable from 0 to 1. Standard errors clustered by country. * denotes significance at 0.10 level (** at 0.05, *** at 0.01).

Table 2.4 reports the results of estimating model (5) in Table 2.2, while

adding interaction terms between particularism and a number of country-specific factors: control of corruption, perceived corruption, corruption diffusion and the overall fraction of particularists. As shown in Section 2.3, the first interaction term is predicted to be negative (i.e., higher corruption deterrence lowers the effect of individual particularism), while the other three interaction terms are expected to be positive. Empirically, all the estimated interaction terms have the expected sign and are statistically significant.

Table 2.4: Determinants of offering a bribe, interactions with particularism

	(1)	(2)	(3)	(4)
Interaction with:				
CoC	-0.09*			
	(0.05)			
CPI		0.41**		
		(0.21)		
Corruption Diffusion			0.02***	
			(0.01)	
Particularism Diffusion				0.00*
				(0.01)
Observations	26429	26429	26330	26330

Note: probit estimates (marginal effects multiplied by 100). Dependent variable: binary variable for having offered a bribe. CPI: Corruption Perception Index. CoC: Control of Corruption. Number of observations: 27807. All specifications include the same set of control variables as in Table 2, column 5. Standard errors clustered by country (multiplied by 100) reported in brackets. * denotes significance at 0.10 level (** at 0.05, *** at 0.01).

In the sociological literature on particularism, Eastern and Southern Europe are generally viewed as particularistic societies, whereas Central and Northern Europe are considered universalist (Hampden-Turner and Trompenaars, 1997).[24] As shown in Section 3.3, the effect of individual-level particularism on the probability of offering a bribe is expected to be stronger where

[24]According to this literature, Central and Northern Europe have been historically shaped through the institutions of the former Austro-Hungarian empire that were more informed by principles of universalism relative to Eastern Europe, where the Ottoman and Communist legacies can be expected to have enhanced particularism (Mungiu-Pippidi, 2005). Moreover, as suggested by Hampden-Turner and Trompenaars (1997), Protestant cultures are more likely to be Universalistic, while predominantly Catholic and Orthodox cultures retain more particularist features.

particularism is more widespread. Table 2.5 reports coefficient estimates for the effect of individual particularism by geographical area (Southern, Eastern, Central and Northern Europe). As expected, the marginal effect of particularism on the probability to offer a bribe is relatively larger in Eastern and Southern Europe (2.45 and 1.34 probability points, respectively), where particularism is more widespread, and smaller in Northern and Central Europe (0.43 and 0.79 probability points, respectively). The difference between Eastern-Southern Europe and Central-Northern Europe is strongly statistically significant (p-value < 0.05).

Table 2.5: Determinants of offering a bribe, by geographical area

	(1) North	(2) Center	(3) East	(4) South
Particularism (d)	0.36**	0.61**	2.03**	0.99**
	(0.16)	(0.28)	(0.94)	(0.31)
Observations	10693	8205	5779	5699

Note: probit estimates (marginal effects multiplied by 100). Dependent variable: binary variable for having offered a bribe. All specifications include the same set of control variables as in Table 2, column 5. North: Denmark, Finland, United Kingdom, Ireland, Iceland, Netherlands, Norway, Sweden South: Spain, Greece, Italy, Portugal, Turkey. East: Czech Republic, Estonia, Poland, Slovenia, Slovakia Center: Switzerland, Germany, Luxembourg, Austria, Belgium, Hungary. Standard errors clustered by country (multiplied by 100) reported in brackets. * denotes significance at 0.10 level (** at 0.05, *** at 0.01).

2.5.2 Accounting for Endogeneity

In order to take into account the possible endogeneity of our key explanatory variable, we estimate the effect of particularism on the probability of offering a bribe by using an Instrumental Variables (IV) estimator. Given the binary nature of the endogenous regressor, we estimate a bivariate marginal effect probit model consisting of two specifications, a reduced form equation for the potentially endogenous dummy variable for particularism and a structural form equation determining the outcome of interest. We included two instruments in the reduced form equation. The first is based on a question

about how important it is for respondents to care for nature and environment. We assume that nature and environment are public goods and that particularist individuals, who have a smaller concern for general interest, care relatively less for them. The second instrument is based on a question regarding whether European unification should go further or whether it has gone too far. As in Uslaner and Conley (2003), we assume that particularist individuals, who are related primarily with people of their own kind, tend to support institutions defending national interests, whereas universalist individuals with looser ties are more likely to support institutions overcoming national interests.

Table 2.6 reports IV estimation results. In the first-stage equation, both instruments have the expected sign, although only the first instrument is significantly related to particularism. The validity of the instruments is not rejected by a Sargan test of over-identifying restrictions ($\chi^2_2 = 0.909$, $p = 0.6346$). The IV estimate of the effect of particularism on the probability of offering a bribe is positive and significant. Indeed, the size of the estimated effect is larger when using IV, suggesting that failing to account for the endogeneity of particularism may lead to underestimate its effect on the propensity to offer a bribe.

As an additional way of addressing the potential endogeneity of particularism, we present results obtained by using Propensity Score (PS) matching estimators (Rosenbaum and Rubin, 1983). These are obtained by using either the nearest neighbor method, which selects the comparison units whose propensity scores are closest to the treated unit in question, or the kernel method, whereby every treated subject is matched with the weighted average of the control subjects, with weights being inversely proportional to the distance between the treated and the control group's propensity scores. In both cases, the computation of the Average Treatment on Treated (ATT) was restricted to the region of common support.[25]

The propensity scores were computed considering variables affecting both

[25]Common support ensures that persons with the same value of the set of observable covariates have a positive probability of being both particularist or non particularist (Heckman et al., 1999). Intuitively, this means that there is enough overlap in the distribution of particularist and non-particularist individuals.

Table 2.6: Determinants of offering a bribe, Instrumental Variables estimation

	Particularism (First Stage)	Offered a bribe (Second Stage)
Particularism		12.55***
Important care nature	-0.34*	
European unification	-0.01	
Gender (d)	2.13***	-1.14***
Age	-0.02	0.01
Age squared	-0.00	0.00
Education	-0.15**	-0.05*
Employed (d)	-0.23	-0.26**
Married (d)	-1.18**	0.58***
Family members	-0.12	0.05
Income	-0.56 *	0.36***
Minority ethnic group	-1.93	0.77**
Immigrant (d)	-3.09***	6.77 *
Son of immigrants (d)	-2.09	3.05
Big city (d)	0.51	-0.23*
Suburbs big city (d)	-0.89	0.69*
Village (d)	2.35***	-0.94***
Friends support	2.03***	-0.72*
Political Interest	0.30	-0.26***
Follow rules	-4.02***	2.36***
Trust	0.17	-0.07**
Trust in legal system	-0.93***	0.49 ***
Trust in public officials	-2.14***	1.00***
Religiosity	-0.37***	0.22***
Risk	-1.55***	0.84***
Male role index	-0.85***	0.44***
Admired	-0.58***	0.36***
Honesty	-4.61***	2.55***
COC	3.76***	-2.39***
CPI	-0.11	-0.31
Observations	27038	27010

Note: IV estimates (marginal effects multiplied by 100). Dependent variable: binary variable for having offered a bribe. (d) indicates discrete change of dummy variable from 0 to 1. Standard errors clustered by country. * denotes significance at 0.10 level (** at 0.05, *** at 0.01).

treatment and outcome (Heckman et al., 1999), fixed over time, and found to be relevant in previous research. Moreover, following Bryson et al. (2002), we preferred to estimate a conservative model with a short list of covariates. More specifically, the first estimated propensity score is based on religiosity, gender and education. The second also adds immigrant status, while the third considers religiosity, gender, education and minority status. The three propensity scores are calculated by restricting the analysis of the balancing property to all treated plus those controls in the region of common support.

Table 2.7 reports PS estimation results. The coefficient for particularism is positive and significant, and virtually unchanged, across all specifications. Overall, we conclude that the findings are robust to the use of alternative estimation techniques that account for the possible endogeneity of our key explanatory variable, and the relationship between particularism and corruption can be given a causal interpretation.

Table 2.7: Determinants of offering a bribe, Propensity score matching estimation

	Nearest Neighbor			Kernel		
	PS 1	PS 2	PS 3	PS 1	PS 2	PS 3
Particularism	1.91***	1.91***	2.01**	2.01**	1.91**	2.02**
	(0.21)	(0.32)	(0.23)	(0.33)	(0.31)	(0.20)

Note: propensity score estimates. Dependent variable: binary variable for having offered a bribe. * denotes significance at 0.10 level (** at 0.05, *** at 0.01). Standard errors (multiplied by 100) reported in brackets.

2.5.3 The Psychological Cost of Bribing

Our theoretical model assumes that particularism reduces the perceived cost of bribing and, through this effect, it increases the probability of offering a bribe. In order to investigate this causal mechanism, we estimate a structural model for the simultaneous relation between particularism, the psychological cost of bribing, and the act of offering a bribe. In the first equation, the probability of being particularist is explained by a set of individual characteristics, including age, gender, education, income, family structure, employment, immigration and minority status. In the second equation, the cost of bribing

is modeled as a function of particularism, the set of exogenous variables described above, Control of Corruption and CPI. In the third equation, offering a bribe is explained by the cost of bribing, particularism, and the same set of individual-level and aggregate control variables.

The results, reported in Table 2.8, support the theoretical hypothesis. Particularism is associated to a significantly lower perceived cost of bribing. Second, through this mechanism, particularism significantly increases the probability of offering a bribe. These results complement the findings in Tavits (2010), indicating that individuals are more likely to engage in bribery when they do not view corruption as wrong. Our results thus contribute to identify the mechanism through which universalism may help to deter corruption.

Table 2.8: Structural Equation Model estimates

	(1) Particularism	(2) Bribe cost	(3) Offered bribe
Particularism		-5.59***	1.59***
		(0.99)	(0.23)
Bribe Cost			-1.35***
			(0.14)
Standard Controls	Yes	Yes	Yes
Observations	26968	26968	26968

Note: Structural Equation Model Estimates. All specifications include the same set of control variables as in Table 2, column 5. Standard errors clustered by country (multiplied by 100) reported in brackets. * denotes significance at 0.10 level (** at 0.05, *** at 0.01).

2.6 Conclusion

Reducing corruption is a key policy objective throughout the world. In order to achieve this objective, it is necessary to have a better understanding of the causes of corruption. Until recently, economists have focused mainly on the macro-level determinants of corruption. Much less is known about the individual-level determinants of corruption and, in particular, about the role

played by individual *cultural* norms.

In this paper, we analyzed theoretically and empirically the effects of the cultural norm of particularism, as opposed to universalism, on collusive bribery. We found that, as predicted by the theory, particularism lowers the perceived cost of corruption and, as a consequence, it has a positive causal effect on the probability of offering bribes. Particularism is also found to have a positive and significant effect on the probability of being asked for a bribe. Our results are robust to the use of alternative definitions of particularism, specifications of the model and econometric techniques to account for the potential endogeneity of particularism. Overall, our results indicate that decreasing particularism can be identified as an effective tool for decreasing corruption. The key question is then how to favor the development of the cultural norm of universalism as opposed to particularism.

Several studies in social psychology show that an individual builds his own social identity on the basis of perceived membership in a social group (Tajfel and Turner, 2004). The intrinsic psychological need for a positive self-image drives individuals to compare their own group with other groups to which they do not belong, giving preferential treatment to members relative to non-members. In this perspective, humans are naturally sectarian and particularism is a feature of human nature that may not be easily changed. Our view, as economists, is slightly different. If particularism can be considered an individual cultural norm, it can be shaped by the social, economic and political environment, albeit in a medium- to long-run time frame (Bisin and Verdier, 2001).

Although it is difficult to indicate specific policy actions aimed at reducing particularism and favouring universalism as a way of combating corruption, education can be identified as the most promising area of intervention. As argued by Glaeser et al. (2007), education enhances social trust thus contributing to overcoming problems of collective action and, in particular, systemic corruption (Uslaner and Rothstein, 2012). Despite the fact that universalism cannot be built by political means (Uslaner, 2013) and that trust is highly persistent over time, since it is transmitted in communities through families (Guiso et al., 2008), several studies have shown that schools, rather than

families, might contribute to increase universalism (Aghion et al., 2010). The relationship between education and universalism, however, should be further explored. There is evidence indicating that schools may affect the building of trust differently depending on their teaching methods (Algan et al., 2013). In addition, it is necessary to better understand how universalism and particularism evolve in societies where they are the predominant cultural norm.

Our findings provide a contribution to the understanding of the determinants of corruption at the individual level. However, they can also be extended to other types of illegal economic behavior. For instance, particularism can be expected to be positively related to tax evasion, since it raises the willingness to free-ride on others. Further research will have to assess the role played by particularism and universalism for other relevant types of anti-social economic behavior.

Appendix A

It is possible to consider cases where it is the public official who asks the citizen for a bribe, rather than the citizen offering a bribe to the public official. The model presented in Section 3 can be easily adapted to describe such transactions. In the first stage, the public official internalizes the possibility for collusion with the citizen and evaluates the costs and benefits of corruption. As above, given that the matching between a citizen and a public official is isolated and random, the probability of successful collusion faced by a public official is given by the fraction of collaborative citizens (β). The public official internalizes this success rate in his decision, and decides whether or not to ask for a bribe. In the second stage, the citizen decides whether or not to accept the public official's proposal for collusion. In the third stage, the bribe amount B is determined by bargaining between the public official and the citizen. The amount of the bribe in equilibrium is therefore the same as in equation (2.9). The citizen is willing to collaborate with the corrupt public official as long as:

$$K - B - qG - \frac{\mu}{\theta}\left[(1-\gamma)H + \gamma L\right] > 0 \tag{2.19}$$

Therefore, given B, the probability that the citizen accepts to collude, is

$$\beta = \theta\left[\frac{K - 2qG - C_p}{(1-\gamma)H + \gamma L}\right]. \tag{2.20}$$

The public official internalizes the probability β that a citizen accepts to offer a bribe and chooses whether to ask for a bribe or not. Since matchings are randomly determined, and the sensitivity of the citizen to the social norm cannot be observed *ex ante*, public officials do not know the exact size of the bribe before the bargain takes place. However, they recognize that collusion can occur if and only if the citizen's sensitivity to the social norm is sufficiently small ($0 < \mu < \mu^*$). Thus, the expected bribery amount $E(B)$,

under the condition $0 < \mu < \mu^*$, is

$$E(B) = \frac{1}{4}K + \frac{3}{4}C_p + \frac{1}{2}qG \tag{2.21}$$

A public official will ask for a bribe as long as

$$E(B) - C_p - qG \geq 0 \tag{2.22}$$

Therefore, the probability that a public official asks for a bribe, that is, the probability that a citizen is asked for a bribe is:

$$\alpha = \int_0^{\varepsilon^*} f(\varepsilon)d\varepsilon = \varepsilon^* = \theta\left[\frac{K - 2qG}{(1-\pi)H + \pi L}\right] \tag{2.23}$$

where $\frac{\partial\alpha}{\partial\pi} > 0$, $\frac{\partial\alpha}{\partial q} < 0$, $\frac{\partial\alpha}{\partial G} < 0$ and $\frac{\partial\alpha}{\partial\theta} > 0$. These partial effects indicate that the higher the probability of being detected, or the higher the size of the fine, the lower is the probability that the citizen is asked for a bribe; in addition, the higher the probability that the public official is particularist, or the higher the perception of overall corruption, the higher is the probability that the citizen is asked for a bribe.

So far, we have assumed that there is uncertainty regarding the type and the psychological cost born by the citizen with whom the public official interacts. This implies that the citizen's particularism is unobservable and, therefore, does not affect the probability of being asked for a bribe. However, it is possible to consider a setting where randomly matched citizens and officials can observe their counterpart's private psychological cost of corruption. In this case, the game becomes a bribery game of complete information (Ryvkin and Serra, 2012). Matchings are still randomly determined, but the sensitivity of the citizen to the social norm and his particularism can be observed ex ante. Public officials therefore do know the exact size of the bribe before the bargaining takes place. Solving the game as above, and keeping the same notation, the probability that a citizen is asked for a bribe is

$$\alpha_p^{perfect} = \left[\frac{\theta(K - 2qG) - \varepsilon L}{(1-\pi)H + \pi L}\right] \tag{2.24}$$

if the citizen is particularist, and

$$\alpha_{np}^{perfect} = \left[\frac{\theta(K - 2qG) - \varepsilon H}{(1-\pi)H + \pi L}\right] \tag{2.25}$$

if the citizen is universalist. Therefore, other things being equal, the probability of being asked for a bribe under complete information is higher for a particularist than for a universalist individual ($\alpha_{np}^{perfect} < \alpha_p^{perfect}$).

Appendix B

Table 2.9: Bribery and particularism, country averages

Countries	Offered bribe	Been Asked for bribe
Austria	0.016	0.055
Belgium	0.008	0.017
Switzerland	0.003	0.013
Czech Republic	0.075	0.132
Germany	0.006	0.020
Denmark	0.007	0.022
Estonia	0.018	0.100
Spain	0.013	0.026
Finland	0.001	0.009
France	0.008	0.014
United Kingdom	0.003	0.012
Greece	0.026	0.118
Hungary	0.013	0.049
Ireland	0.002	0.017
Iceland	0.002	0.014
Italy	0.020	0.053
Luxembourg	0.010	0.036
Netherlands	0.004	0.014
Norway	0.007	0.019
Poland	0.042	0.122
Portugal	0.013	0.036
Sweden	0.007	0.017
Slovenia	0.011	0.033
Slovakia	0.078	0.143
Turkey	0.012	0.067

Source: European Social Survey, 2004-2006. See Section 2.4 for a description of the variables.

Table 2.10: Determinants of offering a bribe, robustness check

	(1) Offered	(2) Offered	(3) Offered	(4) Offered	(5) Offered
Particularism2 (dummy)	0.88***				
	(0.23)				
Particularism3 (dummy)		0.73***			
		(0.14)			
Particularism4			0.43***		
			(0.11)		
Particularism5 (dummy)				0.65**	
				(0.27)	
Particularism6 (dummy)					0.84**
					(0.41)
Observations	26281	26281	26281	26557	26557

Note: probit estimates (marginal effects multiplied by 100). Dependent variable: binary variable for having offered a bribe. All specifications include the same set of control variables as in Table 2.2, column 5. (d) indicates discrete change of dummy variable from 0 to 1. Standard errors (multiplied by 100) clustered by country in parentheses. * denotes significance at 0.10 level (** at 0.05, *** at 0.01).

Bibliography

Aghion, P., Dewatripont, M., Hoxby, C., Mas-Colell, A., and Sapir, A. (2010). The governance and performance of universities: evidence from europe and the us. *Economic Policy*, 25(61):7–59.

Akerlof, G. A. (1980). A theory of social custom, of which unemployment may be one consequence. *The quarterly journal of economics*, pages 749–775.

Alatas, V., Cameron, L., Chaudhuri, A., Erkal, N., and Gangadharan, L. (2009). Gender, culture, and corruption: Insights from an experimental analysis. *Southern Economic Journal*, pages 663–680.

Alesina, A. and Glaeser, E. L. (2004). *Fighting poverty in the US and Europe: A world of difference*. Oxford University Press Oxford.

Algan, Y., Cahuc, P., and Shleifer, A. (2013). Teaching practices and social capital. *American Economic Journal: Applied Economics*, 5(3):189–210.

Baron, D. P. (2010). Morally motivated self-regulation. *The American Economic Review*, 100(4):1299–1329.

Barr, A. and Serra, D. (2010). Corruption and culture: An experimental analysis. *Journal of Public Economics*, 94(11):862–869.

Bisin, A. and Verdier, T. (2001). The economics of cultural transmission and the dynamics of preferences. *Journal of Economic theory*, 97(2):298–319.

Bjørnskov, C. (2006). The multiple facets of social capital. *European journal of political economy*, 22(1):22–40.

Boadway, R., Marceau, N., and Mongrain, S. (2002). Joint tax evasion. *Canadian Journal of Economics*, pages 417–435.

Bowles, R. and Garoupa, N. (1997). Casual police corruption and the economics of crime. *International Review of Law and Economics*, 17(1):75–87.

Brunetti, A. and Weder, B. (2003). A free press is bad news for corruption. *Journal of Public economics*, 87(7):1801–1824.

Bryson, A., Dorsett, R., and Purdon, S. (2002). The use of propensity score matching in the evaluation of active labour market policies.

Cameron, L., Chaudhuri, A., Erkal, N., and Gangadharan, L. (2009). Propensities to engage in and punish corrupt behavior: Experimental evidence from australia, india, indonesia and singapore. *Journal of Public Economics*, 93(7):843–851.

Carey, J. M. and Shugart, M. S. (1995). Incentives to cultivate a personal vote: A rank ordering of electoral formulas. *Electoral studies*, 14(4):417–439.

Chang, J.-j. and Lai, C.-c. (2002). Is the efficiency wage efficient? the social norm and organizational corruption. *The Scandinavian Journal of Economics*, 104(1):27–47.

Chang, J.-j. and Lai, C.-c. (2004). Collaborative tax evasion and social norms: Why deterrence does not work. *Oxford Economic Papers*, 56(2):344–368.

Dancy, J. (1993). *Moral Reasons*. Wiley-Blackwell.

Davis, J. H. and Ruhe, J. A. (2003). Perceptions of country corruption: Antecedents and outcomes. *Journal of Business Ethics*, 43(4):275–288.

De Blasio, G., Scalise, D., and Sestito, P. (2014). Universalism vs. particularism: a round trip from sociology to economics. *Bank of Italy Occasional Paper*, (212).

Dollar, D., Fisman, R., and Gatti, R. (2001). Are women really the "fairer" sex? corruption and women in government. *Journal of Economic Behavior & Organization*, 46(4):423–429.

Dong, B., Dulleck, U., and Torgler, B. (2012). Conditional corruption. *Journal of Economic Psychology*, 33(3):609–627.

Dong, B. and Torgler, B. (2009). Corruption and political interest: empirical evidence at the micro level. *Journal of Interdisciplinary Economics*, 21(3):295–325.

Eckel, C. C. and Grossman, P. J. (2005). Managing diversity by creating team identity. *Journal of Economic Behavior & Organization*, 58(3):371–392.

Fernández, R. and Fogli, A. (2009). Culture: An empirical investigation of beliefs, work, and fertility. *American Economic Journal: Macroeconomics*, 1(1):146–77.

Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika*, 80(1):27–38.

Fisman, R. and Gatti, R. (2002). Decentralization and corruption: evidence across countries. *Journal of Public Economics*, 83(3):325–345.

Fong, C. M. and Luttmer, E. F. (2011). Do fairness and race matter in generosity? evidence from a nationally representative charity experiment. *Journal of Public Economics*, 95(5):372–394.

Glaeser, E. L., Ponzetto, G. A., and Shleifer, A. (2007). Why does democracy need education? *Journal of economic growth*, 12(2):77–99.

Guerrero, M. A. and Rodríguez-Oreggia, E. (2008). On the individual decisions to commit corruption: A methodological complement. *Journal of Economic Behavior & Organization*, 65(2):357–372.

Guiso, L., Sapienza, P., and Zingales, L. (2006). Does culture affect economic outcomes? *Journal of Economic Perspectives*, 20(2):23–48.

Guiso, L., Sapienza, P., and Zingales, L. (2008). Alfred marshall lecture social capital as good culture. *Journal of the European Economic Association*, 6(2-3):295–320.

Gupta, S., Davoodi, H., and Alonso-Terme, R. (2002). Does corruption affect income inequality and poverty? *Economics of governance*, 3(1):23–45.

Hampden-Turner, C. and Trompenaars, F. (1997). *Riding The Waves of Culture: Understanding Diversity in Global Business*. McGraw-Hill.

Harris, D. (2007). Bonding social capital and corruption: a cross-national empirical analysis. Technical report, University of Cambridge, Department of Land Economics.

Heckman, J. J., LaLonde, R. J., and Smith, J. A. (1999). The economics and econometrics of active labor market programs. *Handbook of labor economics*, 3:1865–2097.

Husted, B. W. and y de Estudios, I. T. (1999). Wealth, culture, and corruption. *Journal of International Business Studies*, pages 339–359.

International, T. Transparency international. `https://www.transparency.org/`.

Kant, I. (1788). *The Critique of Practical Reason*. Hackett Publishing Company.

King, G. and Zeng, L. (2001). Logistic regression in rare events data. *Political analysis*, 9(2):137–163.

La Porta, R., Lopez-de Silanes, F., Shleifer, A., and Vishny, R. (1999). The quality of government. *Journal of Law, Economics, and organization*, 15(1):222–279.

Lambsdorff, J. G. (2008). *The Institutional Economics of Corruption and Reform: Theory, Evidence and Policy*. Cambridge University Press.

Lambsdorff, J. G., Taube, M., and Schramm, M. (2004). *The new institutional economics of corruption*. Routledge.

Lee, W.-S. and Guven, C. (2013). Engaging in corruption: The influence of cultural values and contagion effects at the microlevel. *Journal of Economic Psychology*, 39:287–300.

Lipset, S. M. and Lenz, G. S. (2000). Corruption, culture, and markets. *Culture matters: How values shape human progress*, 112:112.

Lumby, J. (2006). International perspectives on leadership and management. *Management in Education*, 20(4):7–10.

Mauro, P. (1995). Corruption and growth. *The quarterly journal of economics*, pages 681–712.

Mauro, P. (1997). *Why worry about corruption?*, volume 6. International Monetary Fund.

Mocan, N. (2008). What determines corruption? international evidence from microdata. *Economic Inquiry*, 46(4):493–510.

Mungiu-Pippidi, A. (2005). Deconstructing balkan particularism: The ambiguous social capital of southeastern europe. *Southeast European and Black Sea Studies*, 5(1):49–68.

Parsons, T. and Shils, E. A. (1951). *Toward a General Theory of Action*. Harvard University Press.

Putnam, R. D. (2001). *Bowling Alone: The Collapse and Revival of American Community*. Touchstone Books by Simon & Schuster.

Rose-Ackerman, S. (2001). Trust, honesty and corruption: Reflection on the state-building process. *European Journal of Sociology*, 42(03):526–570.

Rose-Ackerman, S. (2007). *International handbook on the economics of corruption*. Edward Elgar Publishing.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55.

Ryvkin, D. and Serra, D. (2012). How corruptible are you? bribery under uncertainty. *Journal of Economic Behavior & Organization*, 81(2):466–477.

Serra, D. and Wantchekon, L. (2012). Experimental research on corruption: Introduction and overview. *New advances in experimental research on corruption, research in experimental economics*, 15:1–11.

Spector, B. I. (2005). *Fighting corruption in developing countries: strategies and analysis*. Kumarian Press Bloomfield, CT.

Svensson, J. (2003). Who must pay bribes and how much? evidence from a cross section of firms. *The Quarterly Journal of Economics*, 118(1):207–230.

Swamy, A., Knack, S., Lee, Y., and Azfar, O. (2001). Gender and corruption. *Journal of development economics*, 64(1):25–55.

Tabellini, G. (2010). Culture and institutions: economic development in the regions of europe. *Journal of the European Economic Association*, 8(4):677–716.

Tajfel, H. and Turner, J. C. (2004). The social identity theory of intergroup behavior. In Worchel, S. and Austin, L. W., editors, *Psychology of Intergroup Relations*. Nelson-Hall.

Tavits, M. (2010). Why do people engage in corruption? the case of estonia. *Social forces*, 88(3):1257–1279.

Tonoyan, V. (2003). Corruption and entrepreneurship: Does trust matter? *Emergo. Journal of Transforming Economies and Societies*, 10(3):2–19.

Torgler, B. (2003). Tax morale in transition countries. *Post-communist economies*, 15(3):357–381.

Torgler, B. and Valev, N. T. (2006). Corruption and age. *Journal of Bioeconomics*, 8(2):133–145.

Torgler, B. and Valev, N. T. (2010). Gender and public attitudes toward corruption and tax evasion. *Contemporary Economic Policy*, 28(4):554–568.

Treisman, D. (2000). The causes of corruption: a cross-national study. *Journal of public economics*, 76(3):399–457.

Treisman, D. (2007). What have we learned about the causes of corruption from ten years of cross-national empirical research? *Annu. Rev. Polit. Sci.*, 10:211–244.

Uslaner, E. M. (2002). *The moral foundations of trust*. Cambridge University Press.

Uslaner, E. M. (2013). Trust and corruption revisited: how and why trust and corruption shape each other. *Quality & Quantity*, pages 1–6.

Uslaner, E. M. and Conley, R. S. (2003). Civic engagement and particularized trust the ties that bind people to their ethnic communities. *American Politics Research*, 31(4):331–360.

Uslaner, E. M. and Rothstein, B. (2012). *The Roots of Corruption: Mass education, economic inequality and state building*. Centre for the Study of Public Policy, University of Aberdeen.

Warren, E. M. (2004). What does corruption mean in a democracy? *American journal of political science*, 48(2):328–343.

CHAPTER 3

Extension Services, Production and Welfare: Evidence from a Field Experiment in Ethiopia[1]

[1]This chapter is based on a joint work with Jacopo Bonan (FEEM and LabExpo) and Stefano Pareglio (Catholic University of Brescia). We acknowledge the financial support of Fondazione punto.sud and Fondazione Giangiacomo Feltrinelli. We are very thankful to CIAI staff in Milan and Addis Ababa: Chiara Biffi, Samuela Venturini, Marta Volpi, Yesuf Mohammed. We are indebted to Anna Pasotti and Alessandro Runci for valuable research assistance. We thank Federico Bastia, Fabrizio Alberizzi, Luca Stanca, Mariapia Mendola, Laura Pagani, Stefania Ottone, Claudio Lucifora and Gianluca Femminis for valuable comments. We thank the inhabitants of Ropi and Alemtena for their hospitality.

3.1 Introduction

A large proportion of the world's poor live in rural areas and for 70% of this population agriculture is the main source of income and employment (FAO et al., 2014). In most Sub-Saharan countries agriculture and its associated industries are essential tools to foster a dynamic and inclusive growth and to reduce mass poverty and food insecurity. However, using agriculture as the basis of economic growth requires a productivity revolution in smallholder farming (World Bank, 2007). This revolution has not yet taken place in Sub-Saharan Africa where, despite an increase in agricultural productivity, the agricultural sector lags behind (Evenson, 2003; Krishnan and Patnam, 2014). Empowering rural people by rising their productivity has been and still is an important focus of the agricultural economics research agenda (Udry, 2010; Jack, 2011; Taye, 2013) and, at the same time, it has been the main goal of several development programs (Christiaensen and Demery, 2007).

Low agricultural productivity may be caused by market incompleteness such as credit constraints (Fink et al., 2014; Croppenstedt et al., 2003; Rosenzweig and Wolpin, 1993), imperfect financial and insurance markets (Kazianga and Udry, 2006; Conning and Udry, 2007), weak property rights (Goldstein and Udry, 2008), lack of knowledge and information regarding new technologies, products and methods and their role in enhancing agricultural productivity. Several countries have tried to tackle informational constraints through agricultural extension services, with the goal to include information transfer from the global knowledge base and from local research to farmers, by enabling them to clarify their own goals and possibilities, educating them on how to make better decisions, and stimulating desirable agricultural development (Van der Ban and Hawkins, 1996; Birner et al., 2009).

Despite decades of investment in agricultural extension programs, a general agreement regarding their effectiveness is still lacking and the empirical evidence often gives mixed results, reflecting differences in how the programs are delivered and in the circumstances of beneficiaries (Anderson and Feder, 2007; Aker, 2011). Several scholars show that extension programs have a positive effect on several rural livelihoods including knowledge, skills, productiv-

ity, consumption, and poverty reduction (e.g., Dercon et al. 2009, Godtland et al. 2004, Praneetvatakul and Waibel 2008, Romani 2003, Maffioli et al. 2011, Owens et al. 2003, Davis et al. 2012) while some others find only negligible achievements (e.g., Rivera et al. 2001, Ashraf et al. 2009b, Maffioli et al. 2013). In general, these findings differ according to the technology and context of project implementation, while the estimation of the impact of these programs is often hampered by attribution issues, endogeneity and selection bias. The causal interpretation of the results can be challenging and understanding how new technologies and good practices are *(not)* adopted by farmers is still an open issue (Udry, 2010). For this purpose, the implementation of monitoring and evaluation tools is increasing at all levels, with the aim to provide impact assessments. These can be used by implementers to improve the quality, efficiency and cost-effectiveness of interventions, as the focus is on the results, rather than on the inputs and outputs. Moreover, measuring program impacts raises public awareness and promotes accountability by stakeholders, while, at a more global level, it contributes to building knowledge on what does and does not work to reduce poverty, improve welfare and sustain development.

Assessing the impact of extension services requires to address the attribution issue through the use of identifications which tackle endogeneity problems. Endogeneity may arise from the fact that allocation of extension efforts may not necessarily be random across or within localities. For example, targeting more advantaged (disadvantaged) areas with extension services may respond to different policy aims, but would bias impact estimations upwards (downwards). Moreover, it could be the case that farmers' ability or motivation determines their participation to some activities. This would lead to overestimations of impacts if only more motivated and skilled farmers benefit from extension activities. Few rigorous impact evaluations of extension services in developing countries explicitly tackle the issue of endogeneity and self-selection (Aker, 2011). Different empirical strategies have been adopted, such as quasi-experimental techniques (e.g Godtland et al. 2004), panel data (e.g Maffioli et al. 2011), instrumental variables (e.g Dercon et al. 2009) and randomized experiments (e.g., Ashraf et al. 2009a).

Given the severe limitations in data collection and rigorous research designs, methodological heterodoxy is deemed essential in the study of technological change and agricultural productivity growth (Barrett and Carter, 2010).

In this paper we assess the effect of an agricultural extension project implemented in a village in Ethiopia between 2013 and 2014 and aimed at introducing the cultivation of horticultural gardens along with some more innovative techniques, products and inputs. The project consisted of a series of extension activities including training sessions with the preparation of a show-garden, the distribution of inputs such as vegetable seeds and working tools. Households then received some technical follow-up visits by local development agents. The intervention was targeted to vulnerable households. In accordance with the project's theory of change, our main outcome of interest is the level of adoption of new horticultural products. This specific outcome is crucial to create a new condition where the increased availability of horticultural products may contribute to a more diversified diet and new revenues from sale of products.

The impact of extension programs on agricultural innovation, production and welfare have been thoroughly studied in the literature. Maffioli et al. (2011) evaluate the effectiveness of the Farm Modernization and Development Program in Uruguay and find a positive effect of the program on improved production techniques, such as plantation density and the rate of adoption of certified varieties. Maffioli et al. (2013) find evidence that extension programs increased plantation density. However, they find no evidence of the program's impact on yields for the period under study. Owens et al. (2003) find that, after controlling for innate productivity characteristics and farmers' ability, access to agricultural extension services raises the value of crop production by about 15%. A positive impact of extension services on productivity is also found in Romani (2003) in the Ivory Coast. Dercon et al. (2009) assess the impact of public investments in agricultural extension and road quality on consumption growth and poverty rates in rural Ethiopia. Results suggest that receiving at least one extension visit reduces headcount poverty by 9.8% and increases consumption growth by 7.1%. Davis et al. (2012) estimate the effect of a particular form of agricultural extension service in East

Africa, delivered through farmer field schools, on the economic and production spheres. Results show that farmer field schools have a positive impact on production and income among women, low-literacy, and medium land size farmers. Ashraf et al. (2009b) evaluate the effect of an extension project implemented in Kenya and attempting to support farmers in adopting market export crops by relieving financial and informational constraints. Results suggest that after one year the program led to an increase in the production of export-oriented crops and lower marketing costs; this translated into household income gains for new adopters.

The present work contributes to this literature by exploiting a mixed impact evaluation design combining across-villages comparisons, through difference-in-differences (DID) estimations, with a within village randomized (WVR) control trial design. To this aim, we make use of micro-data collected through surveys administered to 602 households in two time periods (2013 and 2014) to compare the outcomes of three groups: a random selected pool of project beneficiaries in the treated village, a random pool of non-treated farmers in the same area and a pure control group in another village with similar characteristics. We first exploit a DID design to compare changes in the dependent variables of interest between treated and control village across time, conditional on observable characteristics. We also exploit a WVR design and implement an instrumental variable (IV) approach to solve the problem of missing or unknown controls, using the random assignment to treatment and control group as instrument. These exercises, taken together, give us upper and lower-bound estimates of project impacts.

This paper partly contributes also to the literature on the effect of extension services on nutritional outcomes whose link with agricultural interventions has been widely reviewed (see for instance Berti et al. 2004, Girard et al. 2012, Masset et al. 2012, Ruel and Alderman 2013).

Our findings indicate that the project has contributed to production diversification as the number of households growing vegetables increases by about 30 percentage points, corresponding to a 78% increase with respect to the baseline value. We also find a significant increase in revenues from sale of vegetables and, partially, fruits. However, overall, such changes do

not seem to influence total revenues from sales of agricultural products in a relevant way. As far as nutrition is concerned, our results indicate that the intervention did not produce significant changes in the consumption of vegetables and only a marginal increase in fruit uptake. This leads, to an overall irrelevant impact on diet diversification. We provide some possible contextual reason for such results.

The remaining part of this paper is organized as follows: section 3.2 presents the context and the project intervention. Section 3.3 presents sampling strategy and estimation methodology. Section 3.4 reports the results and discussion of the impact evaluation exercise while section 3.5 concludes.

3.2 Context and project intervention

Ethiopia is the second-most populous country in Sub-Saharan Africa with a population of about 90 million (Malik, 2013) and a population growth rate of 2.6% (Michael and Lars Christian, 2013). One of the world's oldest civilizations, Ethiopia is also one of the world's poorest countries with a per capita income of 470 dollars, substantially lower than the regional average (Gross National Income, Atlas Method). The Ethiopian economy has experienced strong growth over the past decade, averaging 10.9% per year in 2004/05 - 2012/13 compared to the regional average of 5.3%. Expansion of the services and agricultural sectors account for most of this growth, while manufacturing sector performance was relatively modest.[2] Agriculture remains mainly traditional and employs more than 80% of the national labor force, accounting for 45% of GDP and 85% of export revenues. Ethiopia is an interesting case study with respect to extension services. In fact, the Ethiopian government has directed for a long time its growth strategy on agriculture by investing close to 1% GDP on these programs (DFID, 2014).

The Siraro district is located in the south-west of the Federal state of Oromia, with a population of around 167,785 inhabitants (CSA, 2014). Siraro is a typical area of the Ethiopian plateau, characterized by an altitude

[2]http://www.worldbank.org/en/country/ethiopia

ranging from 1500 to 2300 meters above sea level. 51.8% of land is arable or cultivable while 1.8% is dedicated to pasture and 4.7% is characterized by forests. The remaining 41.6% is considered swampy, degraded or otherwise unusable. Agriculture is the main activity in the district especially in remote areas. In line with the rest of the country, in the Siraro district there are two crop production seasons referred to as the *meher* (or main) and *belg* (short rain). The *meher* season starts in June and extends to October with the highest concentration in July and August. The *belg* season most often refers to a small but timely rain season, which normally occurs from February to May. The study area includes two villages in the Siraro district, Ropi and Alemtena, located about 10 Km apart. Life expectancy is around 45 years and average age of the population is 16 years. The villages show a substantial shortage of services (e.g. water supply systems, schools, offices, health services, etc). Considering health, three main problems are reported: malaria, malnutrition and water borne diseases. While malaria is reported only during a few months of the year, water borne diseases and malnutrition cases are present permanently. The incidence of malnutrition appears to be particularly relevant for children.

The project first provided eligible farmers with two training sessions which took place in 2013 and 2014, lasting three days and focusing on horticultural process of vegetables and fruits. The vegetables introduced were cabbage, onion, beetroots and carrot along with papaya. Theoretical sessions about the advantages of vegetable and fruit production to improve household welfare were accompanied by the preparation of show-gardens. A specialist from the local health center was invited to illustrate the nutritional properties of vegetables and fruits. Farmers were also trained about the correct use of pesticides, both natural and chemical, to control for major pests and diseases. Some training about soil fertility management, compost making and post-harvest practices was also provided. Participants had the possibility to receive the seeds of the plants discussed in the training sessions and some agricultural tools (hoes and racks). Upon receiving the training and the inputs, farmers were visited by project staff (in collaboration with the local development office and its experts at district level) in order to receive

a technical follow-up on garden maintenance and on the implementation of the techniques explained during the training, such as seed beds preparation, compost and pesticide making, etc. During the visit, a qualitative assessment of the state of the garden and of the implementation of soil fertility management was made by the project staff. The project intervention was designed and implemented by the Italian Center for Children Aid (CIAI), an Italian NGO which has been engaged in cooperation and development activities in Ethiopia since 2003 and which started operating in the municipality of Ropi in 2011.

3.3 Empirical Strategy

3.3.1 Sampling and data

The project was intended to target 250 farmers, 25% of which female-headed households, 8% of model farmers[3], the rest being male-headed, owning at least some land and household size between 5 and 8 members. The first step of the sampling procedure was the selection of beneficiaries farmers in the treated village, following project guidelines. A list of 545 potential beneficiaries satisfying the required criteria was constructed by field project staff, based on local administrative data. Based on the available information, we stratified the sample with respect to per capita land, dividing total land by household size (which we partitioned in tertiles) and gender of the head. We randomly selected from each cell proportionally in order to reach the targeted composition of beneficiaries, for a total of 250 households. The control group in the treated village was randomly selected from the pool of remaining households for each stratum and was drawn with the same proportion employed for the treatment group. The control group was formed by a sample of 121 households.

For the control group in the control village (Alemtena) we started from a

[3]Farmers officially recognized as the most innovative and more productive in the area by government development agents.

list of 425 farmers provided by local authorities, we adopted the same stratification criteria (using the same threshold levels of per capita land) as in the treated village. The sample included a total of 250 households. Unfortunately, the administrative data on which we based the sampling revealed to be slightly different from the actual data collected through surveys. As a consequence, the composition of eligible beneficiaries was slightly different from the initial project objectives.

The baseline survey was conducted in June 2013 before project activities had started, by 10 local enumerators and it included 485 observations: 244 in the treated village and 241 in the control one. Due to financial constraints, only treatment households in the treated village could be surveyed at the baseline.

15 questionnaires could not be included in the analysis and were discarded due to the low quality of data and unreliability of information. The follow-up survey was conducted in November 2014. 481 households visited at the baseline could be successfully tracked, of which 242 in the treated village and 239 in the control one. The very small extent of attrition[4] (less than 1%) and mean comparisons of baseline characteristics between attrited and non-attrited sample lead us to rule out that the final sample of analysis differed significantly from the initial sample of analysis. At the follow-up we also administered the questionnaire to 121 households in the control group in Ropi.

Questionnaires were administered on paper and in local language to the head of the household. In 93% of cases the respondent was the head of household; in the remaining cases the wife (or another key informant) answered the questions, due to the unavailability of the head. Quality and coherence of filled questionnaires were checked by research assistants who also dealt with data-entry using Census and Survey Processing System (CSPro). Data were then cleaned and datasets for analysis were created using Stata software. Descriptive statitistics were produced by computing means and differences in means (using Student's t-tests) across different samples.

[4]The main reason for the inability to track households is due to change of place of residence of the family

3.3.2 Difference-in-Differences (DID)

We first estimate the project impact by using the DID estimator. This method compares the change in the dependent variables of interest in the treated village between two points in time, before and after the project implementation, to the change in the same outcomes in the control area, conditional on observed characteristics. DID relies on the assumption that, conditional on observables, the evolution of the outcomes in treated and control areas would have been the same in the absence of the project, or, equivalently, that any difference in the relevant outcomes between treated and control areas due to unobserved factors is fixed over time (e.g Heckman et al. 1997, Abadie 2005).[5] We estimate the effect of the project on the outcome Y_i using repeated cross-section data. More explicitly, we estimate the following regression with Ordinary Least Squares (OLS)[6]:

$$Y_{it} = \beta_0 + \beta_1 P_i + \beta_2 t + \beta_3 (P_i t) + \gamma x_i + \varepsilon_{it} \tag{3.1}$$

where P_i is a dummy variable equal to one for treatment village (Ropi) and zero for the control one (Alemtena), t is equal to one for post-program and to zero for pre-program. The coefficient attached to the interaction term, β_3, is the parameter of interest and gives the DID estimate of the average effect of the project on outcome Y. β_0 is a constant term, β_1 is the treatment group specific effect which accounts for average permanent differences between treatment and control individuals, β_2 gives the time trend effect common to control and treatment groups, while x_i is a vector of head and household observable baseline characteristics including household's characteristics (households' size, number of children below five in the household, size of the land owned), head of the household characteristics (whether the head is female, whether he/she lives in couple, his/her age, his/her grade of schooling and whether he is a model farmer), participation in the program and in social activities (whether the household has received at least one

[5] This is also known as the parallel trend assumption and means that unobserved characteristics affecting program participation do not vary over time with treatment status.

[6] Heckman and Robb (1985)

visit from development agents, whether the household has benefitted from other support programs, whether the household participates in community based organizations (CBO)[7] and whether the head knows any model farmer). Among the controls we also included the average monthly household's income quintiles and a wealth index computed using Principal Component Analysis (PCA) as suggested by Filmer and Pritchett (2001)[8], by aggregating the information on all assets possessed by the households in a single synthetic index.

While the assignment to the treatment has followed some controlled criteria, there are several reasons why the effective treatment status may be the outcome of some selection process, which goes beyond our control and may hinder the impact evaluation exercise. First of all, among eligible farmers treatment provision was not uniform and homogeneous for all the beneficiaries. A non-random sub-sample received inputs in Summer 2013, a larger sub-sample of eligible farmers received it in Spring 2014, while a relatively small amount of individuals did not receive any project intervention. Moreover, it could be that only more motivated farmers have participated in the training session or may have shown interest in receiving as many inputs as possible. The level of farmers' motivation is a typical unobservable variable which may affect both the likelihood of receiving the intervention and the

[7]CBOs are non-profit groups working at local level to improve life of residents through the provision of services to their members, such as credit and saving opportunities or reciprocal financial support during moments when extraordinary expenditures are required, such as funerals. They can also be focused on human services, natural environment conservation or restoration, and urban environment safety and revitalization.

[8]Formally, the wealth index uses the first principal component of the set of introduced variables. The use of the first principal component yields a wealth index that assigns a larger weight to assets that vary the most across households so that an asset found in all households is given a weight of zero. The first principal component or wealth index can take positive as well as negative values. The categorical variables expressing house facilities such as toilet and water facilities are transformed into ordinal variables and treated as continuous, as suggested by the literature (Vyas and Kumaranayake, 2006). The items considered in the index are: toilet facilities, drinking water facilities, N. of sleeping rooms in the dwelling, hoes, ploughs, granaries, bikes, motorbikes, mobile phones, big-size animals (donkeys, horses, cattle), mid-size animals (goats, sheep), small-size animals (poultry), beds, tables and chairs, kitchen appliances (jerry cans, pots, pans, fans, stoves), energy sources (batteries, generators), other appliances (sewing machines, iron, wheelbarrow, kettle, radio).

final outcomes of interest, for example the adoption of improved techniques and products. In such case we would not be able to disentangle the effect of farmers' motivation from the effect of the project. It could also be that the project staff decided to put more effort in supporting either more skilled and quick learners, as the task would be easier or the more disadvantaged ones (for personal or institutional reasons). Again, it would be hard to distinguish the effect of the selective effort from the real treatment effect. Since we cannot rule out that any of such processes occurred in our study, we evaluate the Intention to Treat Effect (ITT) in (3.1), by considering the overall group of households eligible to receive project activities. The results are therefore being interpreted as the effect of being eligible to participate in the treatment.

Then, in order to check the potential heterogeneous effect of different levels of program take-up we create three categories corresponding to different degrees of program intensity: one corresponds to the uptake of the full package of project activities (participation in training, provision of tools and seeds, at least one technical follow-up visit), the second characterizes farmers who were not exposed to any activity at all while the third is a residual category gathering individuals who were partially exposed to some project activities. Similarly to (3.1), we estimate the effective treatment effect by OLS:

$$Y_{it} = \beta_1 + \beta_2 P_i + \beta_3 t + \beta_4 (AP_i t) + \beta_5 (PP_i t) + \gamma x_i + \varepsilon_{it} \qquad (3.2)$$

where P_i is a dummy variable equal to one for treatment village (Ropi) and zero for the control one (Alemtena); AP_i is a dummy variable equal to one for effective exposure to all project activities, PP_i is equal to one for partial intensity of the project, t is equal to one for post-program and to zero for pre-program. The coefficients attached to the interaction terms, β_4 and β_5, are the parameters of interest and give the DID estimates of the effects of different levels of program intensity with respect to not receiving anything. β_1 is a constant term, β_2 captures group specific effects which accounts for average permanent observable differences between assigned treatment and control individuals, β_3 gives the time trend effect common to control and

treatment groups, while x_i is a vector of head and household observable baseline characteristics as in (3.1).

We cannot interpret, β_4 and β_5 as unbiased average treatment effects of the project since they incorporate both the true treatment effect and the effect of some unobservables. Unfortunately, we are unable to disentangle the two forces (we lack an exogenous instrument), but we can reasonably suppose that the sign of the bias is positive, meaning that we expect that unobservables, such as farmers skills or curiosity, are positively related to both the uptake of the program activities and outcomes. In fact, qualitative discussions with field project staff allow us to rule out that selection of the most disadvantaged has been in place; conversely, farmers demonstrating interest towards the project tended to be more involved by the project staff. Therefore, we are exclusively able to draw conclusions on outcomes which remain unaffected by the project. In other words, on the one hand when estimates are significant, we cannot conclude anything, due to the impossibility to disentangle treatment effect and positive selection bias; on the other hand, when estimates are not significantly different from zero, we can conclude that the project has not impacted the given outcome.

3.3.3 Within village randomization (WVR)

The second evaluation methodology relies on the random assignment of potential project beneficiaries in a treatment and a control group within the village of Ropi. Randomization ensures that the two groups are on average homogeneous along all aspects except the exposure to the treatment, and allows us to attribute causal effects of the project to the differences in outcomes observed. We estimate by OLS the following specification on the sample of Ropi in 2014[9]:

$$Y_i = \beta_0 + \beta_1 P_i + \gamma x_i + \varepsilon_{it} \tag{3.3}$$

P_i is a dummy equal to one for farmers randomly assigned to receive the

[9]As mentioned more extensively in the section on sampling, we do not have baseline data for the control group in Ropi

project and β_1 reports the Intention to Treat Effect (ITT), our parameter of interest. x_i is a vector of observable time-invariant characteristics, including head and household characteristics as in (3.1).

WVR approach bears the risk of violation of the stable unit treatment value assumption (SUTVA), according to which the treatment status of any unit does not affect the potential outcomes of the other units (non-interference). In fact, the control group may directly or indirectly benefit from the exposure to the treatment, for example through the contact and relationship between project beneficiaries and non-beneficiaries which may lead to imitation processes. Moreover, some project activities, such as show gardens, were implemented in public spaces where access was free for all the villagers. We cannot therefore rule out that spillover effects may have influenced non-beneficiaries outcomes. This fact hinders the attribution of effects to the project. However, we can safely assume that spillovers are likely to be positive, meaning that non-beneficiaries may be either directly or indirectly influenced by the project which may change the behavior in the direction of the purposes of the intervention. The evaluation exercise would then tend to underestimate the effects of the programs. In other words, when significant effects are found, those are likely to be underestimated, being even larger in reality, whereas in case of non-significant effects, we cannot conclude anything, due to the impossibility to disentangle the true impact effect from the spillover effect.

We encountered partial compliance issues in the actual project exposure in Ropi, as some eligible farmers did not receive any benefit from the program, whereas some farmers in the control group actually received some benefits which we can observe, through the questionnaire, beyond the unobservable spillover effects. As previously expressed, the mechanism underlying the effective program uptake is non-random and prone to bias. In order to overcome such a problem, we implement an instrumental variable approach (IV) using as instrument the random assignment to the treatment and control groups. We compute the Local Average Treatment Effect (LATE), the

treatment effect on the population of compliers[10], as follows:

$$Y_i = \beta_0 + \beta_1(\hat{EP_i}) + \gamma x_i + \varepsilon_{it} \tag{3.4}$$

where $\hat{EP_i}$ represents the effective treatment exposure and reports the fitted values of the first step regression (using a two stage least squares procedure) in which the effective treatment status is regressed on P_i, the random assignment to treatment variable.

The main purpose of both DID and WVR approaches is to assess average treatment effects on the outcomes of relevance. However, effects may not be necessarily homogeneous throughout the distribution: there might be some classes of individuals who benefit more or less from the intervention. Throughout both methodologies, we explore the existence of heterogeneous effects across some observable characteristics linked to the extent of household vulnerability, by running regressions as in (3.1) and (3.3) on different sub-samples linked to the amount of land owned (above and below the median) and the level of wealth (above and below the median of a wealth index).

3.4 Results

3.4.1 Descriptive statistics

Household characteristics

Table 3.1 show the characteristics of household heads and their families measured in 2013, during the baseline survey.[11] Sample differences in initial conditions between the village of Ropi (treated) and Alemtena (control) are reported in column 4. Lack of significance in sample differences means very

[10] See Angrist et al. (1996) for more details on the methodology

[11] In the case of Ropi control, variables are measured in 2014. In most cases, the variables reported change very slowly in time and are unlikely to be influenced by the project intervention. As a precautionary measure, we asked respondents to date changes of time-variant variables and to repeat comparison of samples on the same time horizon. Results do not vary significantly.

similar initial conditions for treatment and control individuals. Since sample selection of households in Ropi and Alemtena was based on limited and possibly imprecise administrative data, the two samples differ along some characteristics. However, DID approach is such that, as long as differences are assumed constant over time, they do not influence the estimation of treatment effects. Conversely, the use of randomization in the sample selection of individuals in treatment and control groups in Ropi should guarantee that the two groups are identical, on average, across observable and unobservable characteristics, as they only differ in the eligibility to receive the treatment. In the impossibility to show sample differences at the baseline, because we do not have baseline data for Ropi control, we only present differences of variables which are not supposed to change over time. A successful randomization should lead to no significant differences between treatment and control group, however, small sample problems could determine the presence of significant differences. Differences across samples in Ropi are reported in column 5.

Agriculture is the main income generating activity: 90% of the sample works in agriculture, while slightly less than 1% works in the formal sector as public servant, for private companies or NGOs and 2% works in the informal sector, for example running small businesses. Data reveal that even those whose main income generating activity is not agriculture spend some time in growing crops. Income is measured as the sum of all monthly income sources of the household. The average income is 454 BIRR ($ 22) in Ropi and 330 BIRR ($ 16) in Alemtena. The difference varies significantly between sub-samples. The wealth index does not differ significantly across groups. Household heads were asked whether they had received any type of agricultural support provided by NGOs or government in the previous seasons and visits by development agents, governmental extension and agricultural support providers employed by local authorities. With respect to these variables, averages seem to diverge between Ropi and Alemtena (measured at the baseline) and between Ropi and Ropi Control (measured at the follow-up).

The average household is composed of six members of whom at least

one child is less than 5. Land is owned by farmers in almost all our samples. On average, each household own 0.80 hectares. The difference between the amounts of land owned varies significantly among sub-samples: in Ropi, households owns about 0.15 hectares more than in the comparison groups. Female-headed households are about 20% in both Ropi and Alemtena, although they are under-represented in Ropi control. Household heads are about 40 years old and have completed an average of less than 2 years of schooling, however in general in each household at least one member is able to read and to write.

On average, about 10% of our sample is composed of model farmers. Over 80% of the farmers know at least one model farmer. Almost all the households in the sample participate, with some members (normally, but not necessarily the head) in community-based organizations (CBOs). About 19% are members of saving groups, around 78% participate in funeral groups while 20% participate in self-help groups. Significant differences in CBOs involvement arise between Ropi and Alemtena.

Project uptake

According to data collected during the follow-up survey, 93% (226 farmers) of households eligible to receive the treatment declared that they had participated in at least one training session organized within the project. For those not participating, the most common reason cited is linked to the fact that they had not been invited or did not know anything about the project. Among attendants, 82% remembered training and activities on vegetable production, 34% on soil fertility, 23% that on pest control and 22% on diet. 77% of eligible farmers (187 farmers) received at least one tool for agriculture (plough and watering can), 86% of them declared that they had used such tools in the current agricultural season. 82% (199 farmers) declared that they had received seeds [12] but only 52% had actually used them. 85% of eligible farmers visited the show-garden at least once. 89% (213 farmers) received

[12] Among those receiving some seeds, 86% received onion seeds, 96% carrot seeds, 93% beet-root seeds, 94% cabbage seeds

at least one follow-up visit by project staff, on average about 2 visits. 6% (8 farmers out of 121) of the households living in Ropi but not beneficiaries of the project declared that they had participated in some project activities. The garden was visited by 23% of the control sample in Ropi.

Based on this information, the beneficiaries of the program have been classified in three sub-groups according to their degree of project intervention uptake. Beneficiaries who had undertaken all the project activities (training, received at least one tool for agriculture, visited the show garden at least once and had at least one follow-up visit) constitute 64% of the beneficiaries. Beneficiaries who had undertaken none of the project activities are 4% of the sample of beneficiaries. Eligible beneficiaries who had undertaken some are the remaining 32% of the sample.

No households in Alemtena declared that they had been exposed to treatment activity or visited the show garden in Ropi. Only one farmer in the control village declared that he knew project beneficiaries in Ropi. We can therefore exclude any sort of spillover or contamination between the two villages.

Agricultural production and revenues from sale

Table 3.2 provides a characterization of the agricultural sector in the study area, including both Ropi and Alemtena, and presents the main crops cultivated, their main source of seeds, the use of fertilizer, whether the harvest is sold or self-consumed, measured at the baseline. The most commonly cultivated cereal is corn, grown by 98% of households. 30% of households cultivate teff, also known as lovegrass, a particular cereal grown in Ethiopia and Eritrea, where it is used to make enjera, a local substitute for bread. 37% cultivate millet. Cereals' seeds are normally bought at the market and only to a small extent are already owned by families or come from government and NGOs. Much less common is the cultivation of pulses and vegetables, exception made for haricot beans cultivated by almost 50% of the households. Even for pulses and vegetables most of the seeds are purchased at the market, except for carrot seeds that are provided by NGOs in 60% of cases. The use

of fertilizers is quite common for all types of crops. The most sale-oriented crop cultivated is teff (65% of households declared that they had sold all the yields of the previous season) while only 1% of the households declared that they had sold the entire harvest of corn.

Tables 3.3 and 3.4 provide the initial values of outcome variables referring to season 2012, collected at the baseline in Ropi and Alemtena and the values referring to season 2014. In table 3 agricultural enterprises are grouped by main types (cereals, vegetables, fruits, pulses and tubers). Single differences between Ropi and Alemtena samples for each period are shown in columns 3 and 7, while column 9 reports the double difference, which constitutes a first raw evidence of the impact of the project. We also show differences within the treatment and control samples in Ropi, observed in 2014 (column 8).

The initial level of diversification of the production seems relatively low, confirming the picture drawn from table 3.2. On average, farmers grow 2 to 2.5 different crops (significantly more in Ropi than in Alemtena). Preliminary descriptive evidence suggests that the project contributed to increase the variety, which reaches 5 products in the treatment group. Cereals are cultivated by the whole sample, with insignificant differences across samples and time. The share of farmers growing vegetables in 2012 is significantly higher in the Ropi treatment (19%) than in Alemtena (4%). Shares increase in both samples across time, however such growth is much higher in the Ropi treatment group (the share reaches 81%), compared to both Alemtena (31%) and in the Ropi control group (46%). All the products introduced by the project are influencing the total number of products cultivated which increases as does the share of households cultivating vegetables. In particular, cabbage is grown by 70% of the households (40% more than in the control group), carrot by 62% (60% more than the control group) and beet root by 50% (50% more than the control group). The share of farmers growing some fruits in the treatment group also increases significantly, compared to both control groups and across time, although the difference is smaller than in the case of vegetables. About half of the sample also cultivates some pulses and tubers. Despite the presence of sample differences between Ropi and Alemtena at the baseline, no significant changes in such enterprises seem to

occur in time among the two samples.

We compute a measure of sales revenues for the entire agricultural production and for each crop type. Average revenues from sales range from about 1000 to 2000 BIRR ($ 48-96) at the baseline. In general, we observe that the crops having the highest weight on revenues are cereals and pulses while fruits and vegetables have a marginal role. Ropi farmers get more revenues from the sale of products both at the baseline and at the follow-up, compared to the control village. A common time trend leading to higher revenues in the second period in both locations is observed. However, for cereals, vegetables and fruit the increase is higher in the treatment group in Ropi (column 9). Compared to the control sample in Ropi, significantly higher revenues from the sale of vegetables are found in the treatment sample (column 8). We cannot draw any rigorous conclusion on the impact of the intervention from such an exercise which remains mainly descriptive and we postpone the discussion of treatment effects to the next sections.

Food consumption and diet diversification

Table 3.4 shows descriptive statistics on food and nutrition habits. Following international standards, information is obtained from answers to questions concerning the consumption (or lack thereof) of different food items during the week preceding the interview. For the items eaten, the frequency (number of times within the past week), the origin (own harvest vs markets vs donation) and the amount spent are also asked.

It turns out that cereals are the most common food eaten by the households, in their different forms. Cereals or tubers are eaten by all sampled households at least once a week, on average over six times a week; vegetables and milk are consumed at least once a week by 76% and by 20% of households, respectively. Pulses, fruits and meat or fish are consumed by less than 10%. Treated households, after the project intervention, tend to consume cereals, vegetables and meat less frequently. On the contrary, consumption of fruits, oil and fats increases. Even the share of households that declare the consumption of vegetables, fruits, meat or fish at least once a week increases.

However, this information should be jointly considered in order to have a clearer idea of the effect of the project on the households' nutritional status.

A proxy of the level of diet diversification is computed as the simple sum of food types (out of a list of 16 Ethiopian common items included in the questionnaire) which have been eaten at least once in the previous 7 days. Out of 16 types, the average is around 5 at the baseline (slightly higher in the treatment than in the control group), while it grows significantly to 7 and 6.3 at the follow-up in the treatment and control groups, respectively. We also compute the Household Dietary Diversity Score (HDDS) - developed by the Food and Nutrition Technical Assistance Project (FANTA) - which corresponds to the number of different food groups (out of 12), consumed over the past 7 days (Swindale and Bilinsky, 2006). The HDDS food groups are constructed as follows: main staples are disaggregated into two groups (cereals, and roots and tubers), meat, fish and eggs group is disaggregated into its three subgroups; and there is a group for "other foods" , such as condiments, coffee, or tea. The HDDS does not vary significantly between treatment and control group neither at the baseline, nor at the follow-up.

3.4.2 Results

Agricultural production and revenues from sale

Table 3.5 reports the results related to agricultural production and revenues from sale. Within the DID evaluation exercise, column 1 reports ITT as described in (3.1), column 2 and 3 show the coefficient of interest described in (3.2). For what concerns the WVR exercise, column 4 and 5 report the ITT and LATE coefficients described in (3.3) and (3.4), respectively. The project induces a strongly significant increase in the share of farmers cultivating vegetables by over 33 percentage points, confirmed in all specifications. Such a result is likely to be driven by the introduction of the horticultural products promoted and distributed within the project intervention, as witnessed by the increase in farmers growing cabbage (+22 percentage points), carrot (over 50 percentage points), onion (+28 percentage points), beet root

(between 34 and 50 percentage points). The increase in the probability of growing fruits appears lower than vegetables, around 7 percentage points, but it is not confirmed in the within village analysis. In general, the project seems to contribute to a larger product differentiation, as the total number of cultivated products significantly increases by almost 2 items.

Evidence suggests that the number of households which does not sell any product to the market and uses it for self-consumption tends to increase significantly by a range between 8 and 19 percentage points. Concerning sales of products and consequent revenues, we observe a significant increase in the revenues from sale of vegetables, estimated around 150 and 200 BIRR (corresponding to $ 7.2-9.6 per harvest). We find suggestive evidence, not confirmed throughout all specifications, of increases in revenues from sale of cereals and decreases in revenues from sale of pulses and tubers. Overall, such effects lead to no significant changes in total revenues from sale of agricultural products.

The analysis of heterogeneous effects along different levels of owned land and wealth does not point to any significant difference between sub-samples.[13]

It is worth noticing that the five different specifications provide rather similar results in terms of size of coefficients and statistical significance. In particular, in DID specifications no statistical significant difference arises from the comparison of coefficients, meaning that there were no observable relevant differences in outcomes depending on different project intensity. Moreover, ITT and biased ATT estimates are not significantly different from each other, indicating a relatively low level of endogenous unobservable selection.

Estimates from WVR are smaller than DID ones. The former indeed represent lower bounds, given the likely presence of positive spillover effects. ITT and LATE coefficients are relatively similar given the high explanatory power of the instrument in the first step, being around 0.8 throughout the different specifications with a level of significance never above 0.01.[14]

[13]Results are not shown but are available upon request

[14]Results of the first steps are not shown, but are available upon request

Food consumption and diet diversification

Table 3.6 shows the estimated effects of the project on food consumption and diet diversification. There is no evidence of project impact on diet diversification, as measured by the number of food types eaten in the previous seven days and summarized in the HDDS. By looking at the effects for different food categories, we find that no relevant changes occur neither along the extensive nor in the intensive margin consistently across both estimation approaches. DID estimation suggests slight reductions in the frequency of cereals and milk uptake and an increase in the share of people consuming fruits, oil and fats. Such results, however, are not confirmed in the WVR design. It is worth pointing out that, assuming positive within village spillover effects, WVR provides overestimates of the treatment effect. The only significant effect arising from both specifications is the increase in the frequency of fruit consumption, in a range between 0.2 and 0.4 times per week.

3.4.3 Discussion

Results discussion

The project was successful in inducing changes as consequences of the activities proposed and inputs provided. In particular, we find that the number of households growing vegetables increases by about 30 percentage points, i.e. 78% increase compared to the situation at the baseline. This result, in line with the literature (Bushamuka et al., 2005; English and Badcock, 1998; Faber et al., 2002; Laurie and Faber, 2008; Olney et al., 2009), seems to explain a significant increase in revenues from sale of vegetables which is also found by Kumar and Quisumbing (2011). However, overall, such changes do not seem to influence in a relevant way the total revenues from sales of agricultural products, which remain very much anchored to the relevance of cereals and pulses. In terms of revenues from sales, cereals and pulses account for about 88% and 10% of the total, respectively. The role of vegetables seem very marginal, around 1% (the role of fruits is negligible). As a consequence of that, the assessment of project impact on households' economic sphere

(Table 3.7) does not show particularly relevant results. The project does not impact on households' monthly expenditure and on none of its components, namely food, non-food, schooling, healthcare and agricultural expenditures. We do not find conclusive evidence of impacts on the share of people saving and on amounts saved. Other sources of saving and buffer stocks, such as the purchase of cattle or other animals, do not significantly change as a consequence of project activities. We can reasonably rule out that the project caused a short-term raise in disposable income which could be used in expenditure for food, health, schooling or saving. Similar results are also found in the long-run in Kumar and Quisumbing (2011) where a home gardening project is found to have no impact on food and nonfood expenditure, total assets, land owned, and per capita household income.

Households' nutritional sphere does not seem to be affected by project activities. We do not find evidence of improvement in diet diversification, in terms of consumption of vegetables. While we find some indication of changes in the share of households consuming fruits and in the frequency of uptake, the size of the change remains quite small. This result is not surprising in the light of the fact that no storage of vegetable products is available in the area in terms of both processing and conservation. Due to the perishable nature of horticultural products and the impossibility to rely on multiple cycles every year (no irrigation is available and agriculture is mostly rain fed), benefits from horticultural production, both in terms of diet improvement and income generation, can only be considered for a small fraction of the year. This is also found in Hirvonen et al. (2015). Moreover, we cannot rule out that impacts occurred for some particular categories of individuals living in the household, although the overall effect was insignificant, as in Quisumbing and Kumar (2011).

Although our endline survey occurred in a period relatively far from the harvest of the newly introduced horticultural products, our data reveal that the majority of such products has been allocated to households' consumption. Almost 40% and 60% of the households used all vegetables and all fruits harvested exclusively for self-consumption, respectively. The remaining 60% and 40% sold part of their fruits and vegetables to the market, respectively.

This result seems to suggest that, in line with Bushamuka et al. (2005), household beneficiaries could, in principle, generate small income by selling part of their garden production to the market. No households decided to sell on the market the entire production of vegetables and fruits. When looking at the products introduced by the project, the pattern is even more evident. Among 70 and 80% of the households that had actually grown vegetables and fruits promoted by the project used the entire amount harvested for self-consumption. It is likely that such increase in self-consumption of new horticultural products have contributed to short-term improvements in diet diversification, which the timing of our survey prevents from detecting. This result is confirmed in other studies on home-gardening interventions (Faber et al., 2002; Laurie and Faber, 2008; Olney et al., 2009).

Study limitations and robustness checks

Measurement errors are a first source of concerns given that the data collection process is particularly difficult in development contexts, due to, for example, respondents' low levels of education, recall bias, misunderstanding caused by linguistic barriers (given the different local dialects present in the area) or purposely misreporting in the expectation of receiving future support. However, it shall be noted that such issues, if any, should not be distributed in a systematically different manner between treatment and control groups, so that they should not contribute to bias the results. Despite this fact, self-reported values of land extensions do not seem always compatible with the amount of harvest declared, given the regional level of productivity. Thus, we decided not to consider any measure of productivity (like yields per hectare), given the high level of uncertainty in the measure of land extension by each crop. Instead, we rely more on estimates of harvested quantities, as for many products standards of sales and storing exist and are easier to recall (for example corn is stored and sold in sacks of fixed weight). We are unable to rigorously evaluate changes induced by the project in productivity and labor supply within households; we have only indications of an overall increase of the weekly time dedicated to farming. Such an increase is mostly

covered by the household head and wife, when present. Qualitative evidence reveals that all family members take care of the vegetable garden, although this activity is mostly carried out by women, as the garden is normally realized in the proximity of the house. Conversely, men are usually farming cash crops in the larger fields, not necessarily close to the house. As a consequence of these limits in calculating inputs and productivity, we are unable to rigorously assess the profitability of growing horticultural products, in relation to local market prices.

In general, the very early endline survey (less than one year after project implementation) can only capture short-term impacts, leaving out possible longer term effects. Consequently, we should take this into consideration in the evaluation of the chain of changes, whose full development may require several years.

Concerning methodology, both estimation approaches present limits related to the assumptions made or to unobservable factors which may bias the estimates of impacts. A first possible threat to the validity of the DID estimation exercise lies in the possible presence of sample contamination through spillover effects from households in the treatment village to those in the control one. However, since only one household in Alemtena declared that they knew one of the project beneficiaries in Ropi and no household living in Alemtena has visited the show gardens in Ropi, we can rule out spillover effects across villages as a source of identification concern. The second source of concern is related to the DID identifying assumption for which unobservable differences between treatment and control samples are supposed to remain constant over time. In fact, the assumption implies the absence of village-specific shocks, while the very limited number of villages in the treatment and control groups prevents from disentangling village-specific shocks from treatment effects. That said, exposure to shocks does not seem very different for the two villages, given their relative proximity. We look at households exposure to negative shocks occurring between the two rounds of surveys across the two villages, we find that about 50% of households in Ropi experienced some type of shock, while about 35% in Alemtena (the difference is significant at 1%). However, by decomposing the previous variable for

each type of shock, among a list of 13 types, we find that differences across samples are significant only for share of households experiencing flood (5% in Ropi, 0% in Alemtena), heavy rains preventing work (8% vs 0%), increase in price of inputs (8% vs 2%). Overall, the presence of negative shocks to a larger extent in the treatment village may affect the evaluation exercise by underestimating impacts.

In order to better compare treatment and control groups, we repeat the DID exercise, by using propensity score matching. This non-parametric method allows us to balance the two samples along several observable covariates. We construct a propensity score (Rosenbaum and Rubin, 1983) considering variables affecting both treatment and outcome Heckman et al. (1999), fixed over time, and found to be relevant in previous research. Following Bryson et al. (2002), we preferred to estimate a more conservative model with a short list of covariates. More specifically, the estimated propensity score includes indicators of family wealth (the amount of land owned by the household, wealth index and income quintiles), head's characteristics (whether the head is female, whether he/she lives in couple, his/her age, his/her grade of schooling and whether he is a farmer or a model farmer) and indicators of participation in the program and in social activities as described in section 3.2. We implement different matching algorithms, namely the nearest neighbor bias corrected matching estimator (k=1) put forward by Abadie and Imbens (2008), the bootstrapped kernel matching estimator by Becker et al. (2002), the nearest neighbor (k=4) matching with caliper, biweighted kernel matching and radius matching with caliper (Leuven and Sianesi, 2014).[15] In all exercises we impose common support and sample balancing appears satisfactory (the absolute standardized bias[16] is less than 5% for all covariates and the mean propensity score is not different for treated and control individuals in each of six blocks the distribution has been divided in). Results are reported in tables 3.8 and 3.9 in the Appendix and appear

[15]See Caliendo and Kopeinig (2008) for a discussion of trade-off among different matching algorithms

[16]Absolute standardized bias indicates the average percentage difference between treated and control individuals after matching. The literature indicates 5% as the rule of thumb for a satisfactory balancing

in line with estimations obtained with OLS.

Given our small sample and the significant differences between treated and control households in our WVR design, as depicted in column 6 of table 3.1, we re-weight the observations of our control sub-sample in order to perfectly balance covariate distributions in the treated and control groups along the first three sample moments (i.e. mean, variance and skewness), turning to entropy balancing (Hainmueller, 2011; Hainmueller and Xu, 2013). It turns out that none of the results as regards the treatment depend on re-weighting our sample. Results are available upon request.

3.5 Conclusions

We assess the impact of an agricultural extension project implemented in a rural village in Ethiopia between 2013 and 2014 and aimed at introducing the cultivation of horticultural gardens along with innovative techniques, products and inputs. The intervention was targeted to vulnerable households. In accordance with the project's theory of change, our main outcomes of interest is the level of adoption of new horticultural products. As a consequence of adoption, we also assess project impacts on the level of revenues from sale and diet diversification. We find that the project contributes to production diversification as the number of households growing vegetables increases by about 30 percentage points. Overall, such changes do not seem to influence in a relevant way the total revenues from sales of agricultural products, which remain very much anchored to the relevance of cereals and pulses.The results suggest that extension projects similar to the one implemented in the area are an effective strategy for the introduction of vegetables and fruits through home gardens. However, the impact of vegetable sales from small land extensions, as proposed by the project, remains secondary in terms of share of revenues compared to the role of cereals and pulses and does not impact on the total amount of revenues from sales. Although we do not find significant changes in the diet diversification level measured at the time of our survey, we cannot conclude that the program failed to improve the households'

nutritional status. The perishable nature of the products introduced and their seasonality, combined with the relatively small size of land dedicated to these crops, the absence of irrigation and adequate storing technologies are important barriers for the realization of multiple cultivation cycles and the consequent long-term effects on diet diversification. Arguably, better results would have been reached by teaching basic storage and irrigation techniques during the training sessions.

3.6 Tables

Table 3.1: Summary Statistics

	(1)	(2)	(3)	(4)	(5)
	Ropi (R)	Alemtena (A)	Ropi C. (RC)	(R-A)	(R-RC)
HH Characteristics					
HH size	6.46	6.06	6.16	0.40***	-0.30**
N.children less than 5	1.38	1.40	1.13	-0.02	-0.25**
Size of own land	0.77	0.92	0.63	-0.15***	-0.14**
Head characteristics					
Head is female	0.19	0.18	0.12	0.01	-0.072*
Head lives in couple	0.87	0.83	0.82	0.04	-0.04
Age of head	39.97	34.79	38.98	5.18***	-0.99
Head school grade	1.59	1.96	1.61	-0.37	0.02
Head is a model farmer	0.09	0.13	0.06	-0.047*	-0.03
Head is farmer or breeder	0.90	0.88	0.88	0.02	-0.02
Household monthly income	454.36	330.59		123.8***	
Wealth Index	0.07	-0.06		0.13	
Benefit Support Program	0.53	0.80	0.16	-0.27***	-0.37***
At least one visit from DA	0.90	0.93	1.00	-0.03	0.095***
Participation in CBO	0.88	0.70	0.91	0.18***	0.03
Know model farmer	0.81	0.91	0.75	-0.098***	-0.06
Observations	242	239	121	481	363

Note: * denotes significance at 0.10 level (** at 0.05, *** at 0.01).

Table 3.2: Agricultural production in season 2012, descriptive statistics, all sample

		Seed origin					Sale of harvest		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Crop	Share of HH cultivating	Own	Market	Government or NGO	Share of HH using fertilizer	Avg sale price (Birr/quintal)	All	Partial	None
Cereals									
Corn	0.981	0.077	0.772	0.143	0.815	425	0.015	0.340	0.621
Millet	0.981	0.077	0.772	0.143	0.815	425	0.015	0.340	0.621
Sorghum	0.079	0.158	0.842	0.000	0.711	499	0.105	0.026	0.842
Teff	0.301	0.063	0.924	0.014	0.951	1122	0.653	0.222	0.111
Pulses									
Haricot beans	0.470	0.120	0.876	0.004	0.711	517	0.418	0.253	0.320
Vegetables									
Cabbage	0.096	0.435	0.543	0.022	0.761	148	0.130	0.152	0.652
Carrot	0.010	0.000	0.400	0.600	1.000	340	0.200	0.200	0.600
Pepper	0.015	0.429	0.571	0.000	0.714	272	0.286	0.143	0.429

Note: only crops produced by at least 1% of households are reported.

Table 3.3: Agricultural production, sale and value. Means at the baseline (2012), follow up (2014), simple and double differences

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Ropi (R0)	Alemtena (A0)	Difference (R0-A0)	Ropi (R1)	Alemtena (A1)	Ropi C. (RC1)	Difference (R1-A1)	Difference (R1-RC1)	D-D (R1-A1)-(R0-A0)
Agricultural production									
Number of products cultivated	2.48	2.19	0.291***	5.27	3.02	3.40	2.248***	1.86***	1.957***
Cereals are grown	0.98	0.98	0.00	0.99	0.97	0.96	0.02	0.029*	0.02
Vegetables are grown	0.19	0.04	0.152***	0.81	0.31	0.46	0.500***	0.34***	0.348***
Fruits are grown	0.02	0.00	0.021**	0.10	0.01	0.04	0.095***	0.062**	0.074***
Pulses and tubers are grown	0.52	0.42	0.102**	0.56	0.56	0.50	0.00	0.07	-0.10
Total revenues from sale of									
Total crop sale	2144.76	944.08	1200.677***	2942.18	1176.30	2932.77	1765.882***	9.41	565.206***
Cereals	1748.77	769.62	979.145***	2182.70	366.82	2439.79	1815.877***	-257.10	836.731***
Vegetables	25.83	2.30	23.525***	214.24	17.60	49.34	196.645***	164.9***	173.119***
Fruits	0.00	0.00	0.000***	3.22	0.00	0.00	3.223*	3.22	3.223*
Pulses and tubers	369.41	172.16	197.254***	539.96	791.88	443.64	-251.928**	96.30	-449.183**

Note: * denotes significance at 0.10 level (** at 0.05, *** at 0.01).

Table 3.4: Food and nutrition. Means at the baseline, follow up and differences

		Baseline			Follow-up					
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Ropi (R0)	Alemtena (A0)	Difference (R0-A0)	Ropi (R1)	Alemtena (A1)	Ropi C. (RC1)	Difference (R1-A1)	Difference (R1-C1)	D-D (R1-A1)-(R0-A0)
Cereals or tubers	at least once	1.00	1.00	0.00	1.00	1.00	1.00	0.00	0.00	0.00
	n. of times	6.53	6.05	0.479***	6.76	6.92	6.68	-0.156***	0.09	-0.635***
Pulses	at least once	0.06	0.04	0.02	0.14	0.13	0.10	0.00	0.04	-0.02
	n. of times	0.17	0.14	0.02	0.32	0.39	0.20	-0.07	0.12	-0.09
Vegetables	at least once	0.76	0.87	-0.110***	0.91	0.97	0.88	-0.062***	0.03	0.048***
	n. of times	3.75	4.46	-0.708***	4.29	5.66	3.98	-1.364***	0.31	-0.655***
Fruits	at least once	0.05	0.03	0.02	0.18	0.03	0.08	0.157***	0.11***	0.136***
	n. of times	0.12	0.11	0.02	0.38	0.08	0.13	0.297***	0.23***	0.281***
Meat or fish	at least once	0.07	0.02	0.058***	0.18	0.05	0.21	0.123***	-0.03	0.066***
	n. of times	0.28	0.05	0.235***	0.29	0.11	0.25	0.180***	0.04	-0.054***
Milk	at least once	0.20	0.14	0.060*	0.36	0.39	0.24	-0.03	0.12**	-0.09
	n. of times	0.67	0.50	0.17	1.14	1.42	0.79	-0.28	0.35*	-0.45
Oil or fats	at least once	0.65	0.87	-0.222***	0.91	0.96	0.87	-0.053**	0.04	0.168**
	n. of times	2.57	3.37	-0.806***	4.19	4.80	3.80	-0.605***	0.39	0.201***
# of food types, past 7 days		5.25	4.90	0.348***	7.02	6.34	6.67	0.678***	0.35*	0.329***
HDDS		4.41	4.44	-0.03	5.34	5.25	4.90	0.10	0.44***	0.13

Note: * denotes significance at 0.10 level (** at 0.05, *** at 0.01).

Table 3.5: Project effects on agricultural production DID (Ropi vs Alemtena) and within village randomization (Ropi beneficiaries vs Ropi control)

	DID			WVR	
	ITT	ATT*		ITT	LATE
		Full	Partial		
	(1)	(2)	(3)	(4)	(5)
Agricultural production					
Cereals are grown	0.021	0.013	0.024	0.023	0.027
	0.017	0.017	0.020	0.017	0.021
Vegetables are grown	0.348***	0.331***	0.418***	0.338**	0.386***
	0.048	0.052	0.054	0.057	0.062
Fruits are grown	0.074***	0.081***	0.072**	0.024	0.027
	0.022	0.027	0.036	0.024	0.041
Pulses and tubers are grown	-0.101	-0.084	0.049	0.018	0.021
	0.062	0.066	0.075	0.063	0.072
Total number of cultivated products	1.957***	1.913***	2.327***	1.581***	1.806***
	0.191	0.226	0.252	0.207	0.296
Products introduced by the project					
Cabbage is grown	0.261***	0.220***	0.355***	0.219***	0.265***
	0.049	0.055	0.061	0.062	0.070
Carrot is grown	0.587***	0.603***	0.598***	0.480***	0.561***
	0.033	0.041	0.055	0.049	0.062
Onion is grown	0.285***	0.307***	0.283***	0.189***	0.243***
	0.030	0.037	0.052	0.040	0.057
Beet root is grown	0.496***	0.507***	0.498***	0.342***	0.405***
	0.032	0.042	0.059	0.054	0.065
Papaya is grown	0.050***	0.061***	0.034	0.030**	0.033
	0.018	0.023	0.026	0.014	0.031
Total revenues from sale of					
Entire production for self-consumption	0.191***	0.157***	0.197***	0.076**	0.087**
	0.047	0.046	0.054	0.030	0.036
Cereals	836.73**	1306.22***	933.69*	-591.34	-675.52
	368.332	431.138	509.421	610.745	586.917
Vegetables	173.11***	171.58***	220.14***	145.51***	166.23***
	35.030	41.927	61.208	39.826	62.164
Fruits	3.223*	3.832*	1.952	0.958	1.094
	1.741	2.313	1.944	0.954	3.297
Pulses and tubers	-449.18***	-251.37*	-418.02***	45.17	51.60
	128.007	145.625	126.842	116.524	144.583
Total revenues from crop sale	565.206	1233.047**	737.387	-397.719	-454.335
	413.917	483.916	552.985	638.953	626.535
Observations	962	962	962	361	361

Note: the effects of the program on the variable reported in rows are shown with robust standard errors reported under the coefficient. DID regressions include the set of controls as described in section 3.1. * denotes significance at 0.10 level (** at 0.05, *** at 0.01).

Table 3.6: Project effects on food consumption and diet diversification. DID (Ropi vs Alemtena) and within village randomization (Ropi beneficiaries vs Ropi control)

		DID			WVR	
		ITT	ATT*		ITT	LATE
			Full	Partial		
		(1)	(2)	(3)	(4)	(5)
N. of food types, past 7 days		0.329*	0.335*	0.410	0.060	0.066
		0.192	0.196	0.256	0.241	0.252
HDDS		0.126	0.207	0.111	0.235	0.257
		0.128	0.128	0.174	0.149	0.163
		0.044	0.047	0.052	0.039	0.053
Food categories						
Cereals	at least once	0.000	0.000	0.000	0.000	0.000
		0.000	0.000	0.000	0.000	0.000
	n. of times	-0.635***	-0.596***	-0.608***	-0.022	-0.025
		0.125	0.120	0.149	0.102	0.114
Pulses	at least once	-0.022	-0.010	-0.056	-0.017	-0.023
		0.036	0.040	0.044	0.042	0.048
	n. of times	-0.086	-0.116	-0.234*	0.039	0.045
		0.118	0.136	0.132	0.095	0.143
Vegetables	at least once	0.048	0.027	0.095**	0.013	0.015
		0.041	0.043	0.044	0.038	0.045
	n. of times	-0.655**	-0.622**	-0.375	-0.097	-0.144
		0.274	0.283	0.303	0.236	0.298
Fruits	at least once	0.136***	0.122***	0.223***	0.058	0.069
		0.032	0.035	0.052	0.043	0.052
	n. of times	0.281***	0.231**	0.449***	0.193**	0.178
		0.096	0.097	0.131	0.092	0.115
Meat	at least once	0.066*	0.083**	0.069	-0.017	-0.020
		0.034	0.039	0.049	0.047	0.049
	n. of times	-0.054	-0.058	-0.159	0.026	0.029
		0.101	0.124	0.112	0.067	0.109
Milk	at least once	-0.086	-0.044	-0.076	0.052	0.048
		0.053	0.056	0.065	0.058	0.066
	n. of times	-0.445**	-0.267	-0.405*	0.146	0.116
		0.207	0.218	0.236	0.177	0.240
Oils and fats	at least once	0.168***	0.166***	0.180***	-0.007	-0.004
		0.043	0.044	0.047	0.040	0.045
	n. of times	0.201	0.441*	0.384	-0.082	-0.123
		0.257	0.265	0.305	0.240	0.311
Observations		962	962	962	361	361

Note: the effects of the program on the variable reported in rows are shown with robust standard errors reported under the coefficient. DID regressions include the set of controls as described in section 3.1. * denotes significance at 0.10 level (** at 0.05, *** at 0.01).

Table 3.7: Project effects on expenditure and savings DID (Ropi vs Alemtena) and within village randomization (Ropi beneficiaries vs Ropi control)

	DID			WVR	
	ITT	ATT*		ITT	LATE
		Full	Partial		
	(1)	(2)	(3)	(4)	(5)
Total monthly expenditure	70.071	118.711	74.140	62.995	76.644
	92.784	93.890	98.407	70.288	79.120
Monthly food expenditure	121.632	154.359**	140.304*	-1.859	6.130
	74.713	77.067	74.438	51.525	56.956
Total monthly non-food expenditure	-27.681	-22.679	-49.332	1.451	2.616
	30.115	29.631	33.707	30.395	34.311
Monthly school expenditure	-4.003	-3.062	-1.428	-0.869	-0.830
	3.293	2.955	3.571	1.777	2.132
Monthly healthcare expenditure	1.307	6.259	3.386	14.695*	17.508
	8.747	9.480	9.592	8.127	11.576
Monthly agriculture expenditure	-16.802*	-15.103	-17.007	14.842	15.880
	9.493	10.455	12.362	9.586	11.757
Other expenditures	-3.449	0.217	-1.839	30.904***	31.192**
	12.772	13.282	16.199	10.909	14.848
Cash saving	0.022	-0.038	-0.017	0.104**	0.129**
	0.048	0.053	0.066	0.051	0.063
Amount saved	73.579	81.810	81.121	-132.094	-150.898
	56.773	63.538	98.003	147.157	152.323
Observations	962	962	962	361	361

Note: the effects of the program on the variable reported in rows are shown with robust standard errors reported under the coefficient. DID regressions include the set of controls as described in section 3.1. * denotes significance at 0.10 level (** at 0.05, *** at 0.01).

3.7 Appendix

Table 3.8: Project effect on agricultural production and household expenditures, Propensity Score Matching Estimation

	(1)	(2)	(3)	(4)	(5)
Outcome	Nearest Neighbor	Kernel	Caliper	Biweighted kernel	Radius
Agricultural Production					
Cereals are grown	0.021	0.021	0.004	0.021	0.021
	0.014	0.013	0.009	0.014	0.014
Vegetables are grown	0.348***	0.348***	0.866***	0.348***	0.348***
	0.048	0.047	0.252	0.048	0.048
Fruits are grown	0.074***	0.074***	0.083***	0.074***	0.074***
	0.02	0.02	0.02	0.02	0.02
Pulses are grown	-0.101*	-0.101*	-0.709***	-0.101*	-0.101*
	0.055	0.064	0.254	0.061	0.061
# of cultivated products	1.957***	1.957***	2.035***	1.957***	1.957***
	0.176	0.188	0.648	0.182	0.182
Products introduced by the project					
Papaya is grown	0.05***	0.05***	0.054***	0.05	0.05***
	0.018	0.018	0.017	0.017	0.017
Cabbage is grown	0.261***	0.261***	0.521***	0.261***	0.261***
	0.050	0.051	0.037	0.049	0.049
Carrot is grown	0.587***	0.587***	0.595***	0.587***	0.587***
	0.032	0.033	0.032	0.033	0.033
Onion is grown	0.285***	0.285***	0.285***	0.285***	0.285***
	0.030	0.029	0.029	0.029	0.029
Beet root is grown	0.496***	0.496***	0.496***	0.496***	0.496***
	0.034	0.030	0.032	0.032	0.032
Total revenues from sale of					
Cereals	836.731**	836.731**	933.928*	836.731**	836.731**
	419.796	396.388	493.135	408.045	408.045
Vegetables	173.119***	173.119***	300.917**	173.119***	173.119***
	38.564	33.712	116.987	34.098	34.098
Fruits	3.223*	3.223*	3.223*	3.223*	3.223*
	1.647	1.738	1.737	1.737	1.737
Pulses	-449.183***	-449.183***	-804.455	-449.183***	-449.183***
	123.811	122.351	827.649	127.011	127.011
Crop sale	565.206	565.206	434.927	565.206	565.206
	458.773	504.603	846.141	458.065	458.065

Note: the effects of the program on the variable reported in rows are shown with robust standard errors reported under the coefficient. DID regressions include the set of controls as described in section 3.1. * denotes significance at 0.10 level (** at 0.05, *** at 0.01).

Table 3.9: Project effect on nutrition, Propensity Score Matching Estimation

	(1)	(2)	(3)	(4)	(5)
Outcome	Nearest Neighbor	Kernel	Caliper	Biweighted kernel	Radius
Nutrition					
# of food types, past 7d	0.329*	0.329*	0.773*	0.329*	0.329*
	0.183	0.198	0.432	0.186	0.186
HDDS	0.126	0.126	0.438	0.126	0.126
	0.124	0.119	0.304	0.128	0.128
At least once per week					
Eat cereals and tubers	0***	0***	0***	0***	0***
	0	0	0	0	0
n. of times eaten					
cereals and tubers	-0.635***	-0.635***	0.231***	-0.635***	-0.635***
	0.124	0.124	0.081	0.127	0.127
Eat pulses	-0.022	-0.022	0.074***	-0.022	-0.022
	0.037	0.036	0.026	0.035	0.035
n. of times eaten pulses	-0.09	-0.09	0.153*	-0.09	-0.09
	0.11	0.139	0.078	0.117	0.117
Eat vegetables	0.048	0.048	0.145***	0.048	0.048
	0.04	0.04	0.033	0.04	0.04
n. of times eaten vegetables	-0.655**	-0.655**	-0.455	-0.655**	-0.655**
	0.286	0.727	1.1	0.27	0.27
Eat fruits	0.136***	0.136***	0.136***	0.136***	0.136***
	0.035	0.034	0.029	0.032	0.032
n. of times eaten fruits	0.265***	0.265***	0.24***	0.265***	0.265***
	0.087	0.094	0.073	0.095	0.095
Eat meat/egg	0.066**	0.066**	0.103***	0.066**	0.066**
	0.034	0.034	0.029	0.033	0.033
n. of times eaten meat/egg	-0.054	-0.054	0.008	-0.054	-0.054
	0.091	0.122	0.092	0.101	0.101
Eat milk	-0.086*	-0.086*	-0.339	-0.086*	-0.086*
	0.053	0.052	0.291	0.051	0.051
n. of times eaten milk	-0.445**	-0.445**	-0.029	-0.445**	-0.445**
	0.194	0.181	0.317	0.193	0.193
Eat oil/fats	0.168***	0.168***	0.26***	0.168***	0.168***
	0.046	0.041	0.034	0.043	0.043
n. of times eaten oil/fats	0.201	0.201	0.628	0.201	0.201
	0.272	0.26	1.241	0.26	0.26

Note: the effects of the program on the variable reported in rows are shown with robust standard errors reported under the coefficient. DID regressions include the set of controls as described in section 3.1. * denotes significance at 0.10 level (** at 0.05, *** at 0.01).

Bibliography

Abadie, A. (2005). Semiparametric difference-in-differences estimators. *The Review of Economic Studies*, 72(1):1–19.

Abadie, A. and Imbens, G. W. (2008). On the failure of the bootstrap for matching estimators. *Econometrica*, 76(6):1537–1557.

Aker, J. C. (2011). Dial a for agriculture: a review of information and communication technologies for agricultural extension in developing countries. *Agricultural Economics*, 42(6):631–647.

Anderson, J. R. and Feder, G. (2007). *Agricultural Extension*, volume 3 of *Handbook of Agricultural Economics*, chapter 44, pages 2343–2378. Elsevier.

Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American statistical Association*, 91(434):444–455.

Ashraf, N., Berry, J., and Shapiro, J. M. (2009a). Can higher prices stimulate product use? evidence from a field experiment in Zambia. Working paper, National Bureau of Economic Research.

Ashraf, N., Giné, X., and Karlan, D. (2009b). Finding missing markets (and a disturbing epilogue): Evidence from an export crop adoption and marketing intervention in Kenya. *American Journal of Agricultural Economics*, 91(4):973–990.

Barrett, C. B. and Carter, M. R. (2010). The power and pitfalls of experiments in development economics: Some non-random reflections. *Applied Economic Perspectives and Policy*, 32(4):515–548.

Becker, S. O., Ichino, A., et al. (2002). Estimation of average treatment effects based on propensity scores. *The stata journal*, 2(4):358–377.

Berti, P. R., Krasevec, J., and FitzGerald, S. (2004). A review of the effectiveness of agriculture interventions in improving nutrition outcomes. *Public Health Nutrition*, 7:599–609.

Birner, R., Davis, K., Pender, J., Nkonya, E., Anandajayasekeram, P., Ekboir, J., Mbabu, A., Spielman, D. J., Horna, D., Benin, S., and Cohen, M. (2009). From best practice to best fit: A framework for designing and analyzing pluralistic agricultural advisory services worldwide. *The Journal of Agricultural Education and Extension*, 15(4):341–355.

Bryson, A., Dorsett, R., and Purdon, S. (2002). The use of propensity score matching in the evaluation of active labour market policies. Working paper.

Bushamuka, V. N., de Pee, S., Talukder, A., Kiess, L., Panagides, D., Taher, A., and Bloem, M. (2005). Impact of a homestead gardening program on household food security and empowerment of women in Bangladesh. *Food & Nutrition Bulletin*, 26(1):17–25.

Caliendo, M. and Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of economic surveys*, 22(1):31–72.

Christiaensen, L. and Demery, L. (2007). *Down to Earth*. World Bank.

Conning, J. and Udry, C. (2007). *Chapter 56 Rural Financial Markets in Developing Countries*, volume 3 of *Handbook of Agricultural Economics*, pages 2857 – 2908. Elsevier.

Croppenstedt, A., Demeke, M., and Meschi, M. M. (2003). Technology adoption in the presence of constraints: the case of fertilizer demand in Ethiopia. *Review of Development Economics*, 7(1):58–70.

CSA (2014). *National report, Ethiopian Central Statistical Agency.*

Davis, K., Nkonya, E., Kato, E., Mekonnen, D., Odendo, M., Miiro, R., and Nkuba, J. (2012). Impact of Farmer Field Schools on Agricultural Productivity and Poverty in East Africa. *World Development*, 40(2):402–413.

Dercon, S., Gilligan, D. O., Hoddinott, J., and Woldehanna, T. (2009). The impact of agricultural extension and roads on poverty and consumption growth in fifteen Ethiopian villages. *American Journal of Agricultural Economics*, 91(4):1007–1021.

DFID (2014). *Can agriculture interventions promote nutrition?* DFID Evidence Papers.

English, R. and Badcock, J. (1998). A community nutrition project in Viet Nam: Effects on child morbidity. *Food Nutrition and Agriculture*, pages 15–21.

Evenson, R. (2003). *Crop variety improvement and its effect on productivity the impact of international agricultural research.* CABI Pub.

Faber, M., Venter, S. L., and Benade, A. (2002). Increased vitamin a intake in children aged 2-5 years through targeted home-gardens in a rural South African community. *Public health nutrition*, 5(01):11–16.

FAO, IFAD, and WFP (2014). *State of food insecurity in the world 2014: strengthening the enabling environment for food security and nutrition.* Food and Agriculture Organization of the United Nations, Rome.

Filmer, D. and Pritchett, L. H. (2001). Estimating wealth effects without expenditure data or tears: An application to educational enrollments in states of India. *Demography*, 38(1):115–132.

Fink, G., Jack, B. K., and Masiye, F. (2014). *National Bureau of Economic Research*.

Girard, A. W., Self, J. L., McAuliffe, C., and Olude, O. (2012). The effects of household food production strategies on the health and nutrition outcomes of women and young children: A systematic review. *Paediatric and Perinatal Epidemiology*, 26:205–222.

Godtland, E. M., Sadoulet, E., Janvry, A. D., Murgai, R., and Ortiz, O. (2004). The impact of farmer field schools on knowledge and productivity: A study of potato farmers in the Peruvian Andes. *Economic Development and Cultural Change*, 53:63–92.

Goldstein, M. and Udry, C. (2008). The profits of power: Land rights and agricultural investment in Ghana. *Journal of Political Economy*, 116(6):pp. 981–1022.

Hainmueller, J. (2011). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, page mpr025.

Hainmueller, J. and Xu, Y. (2013). Ebalance: A stata package for entropy balancing. *Journal of Statistical Software*, 54(7).

Heckman, J. J., Ichimura, H., and Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The Review of Economic Studies*, 64(4):pp. 605–654.

Heckman, J. J., LaLonde, R. J., and Smith, J. A. (1999). The economics and econometrics of active labor market programs. *Handbook of labor economics*, 3:1865–2097.

Heckman, J. J. and Robb, R. (1985). Alternative methods for evaluating the impact of interventions: An overview. *Journal of econometrics*, 30(1):239–267.

Hirvonen, K., Taffesse, A. S., and Worku, I. (2015). Seasonality and household diets in Ethiopia:. Technical report.

Jack, B. K. (2011). White paper prepared for the agricultural technology adoption initiative. Working paper.

Kazianga, H. and Udry, C. (2006). Consumption smoothing? livestock, insurance and drought in rural Burkina Faso. *Journal of Development Economics*, 79(2):413 – 446.

Krishnan, P. and Patnam, M. (2014). Neighbors and extension agents in ethiopia: Who matters more for technology adoption? *American Journal of Agricultural Economics*, 96(1):308–327.

Kumar, N. and Quisumbing, A. R. (2011). Access, adoption, and diffusion: understanding the long-term impacts of improved vegetable and fish technologies in Bangladesh. *Journal of Development Effectiveness*, 3(2):193–219.

Laurie, S. M. and Faber, M. (2008). Integrated community-based growth monitoring and vegetable gardens focusing on crops rich in β-carotene: Project evaluation in a rural community in the Eastern Cape, South Africa. *Journal of the Science of Food and Agriculture*, 88(12):2093–2101.

Leuven, E. and Sianesi, B. (2014). Psmatch2: Stata module to perform full mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. *Statistical Software Components*.

Maffioli, A., Ubfal, D., Baré, G. V., and Cerdán-Infantes, P. (2011). Extension services, product quality and yields: the case of grapes in Argentina. *Agricultural Economics*, 42(6).

Maffioli, A., Ubfal, D., Vazquez-Bare, G., and Cerdan-Infantes, P. (2013). Improving technology adoption in agriculture through extension services: evidence from Uruguay. *Journal of Development Effectiveness*, 5(1):64–81.

Malik, K. (2013). *Human development report 2013: the rise of the South. Human progress in a diverse world*. United Nations Development Programme.

Masset, E., Haddad, L., Cornelius, A., and Isaza-Castro, J. (2012). Effectiveness of agricultural interventions that aim to improve nutritional status of children: systematic review. *BMJ*, 344.

Michael, G. and Lars Christian, M. (2013). Ethiopia - second economic update: laying the foundation for achieving middle income status. Working paper, World Bank.

Olney, D. K., Talukder, A., Iannotti, L. L., Ruel, M. T., and Quinn, V. (2009). Assessing impact and impact pathways of a homestead food production program on household and child nutrition in Cambodia. *Food & Nutrition Bulletin*, 30(4):355–369.

Owens, T., Hoddinott, J., and Kinsey, B. (2003). The impact of agricultural extension on farm production in resettlement areas of Zimbabwe. *Economic Development and Cultural Change*, 51(2).

Praneetvatakul, S. and Waibel, H. (2008). A panel data model for the assessment of farmer field schools in thailand. *Agriculture in Developing Countries-Technology Issues. Sage Publishers, London*, pages 137–151.

Quisumbing, A. R. and Kumar, N. (2011). Does social capital build women's assets? the long-term impacts of group-based and individual dissemination of agricultural technology in Bangladesh. *Journal of Development Effectiveness*, 3(2):220–242.

Rivera, W. M., Qamar, M. K., Van Crowder, L., et al. (2001). *Agricultural and rural extension worldwide: options for institutional reform in the developing countries.* FAO, Rome.

Romani, M. (2003). *The impact of extension services in times of crisis: Côte d'Ivoire (1997-2000)*. University of Oxford, Centre for the Study of African Economies.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55.

Rosenzweig, M. R. and Wolpin, K. I. (1993). Credit market constraints, consumption smoothing, and the accumulation of durable production assets in low-income countries: Investment in bullocks in India. *Journal of Political Economy*, 101(2):223–44.

Ruel, M. T. and Alderman, H. (2013). Nutrition-sensitive interventions and programmes: how can they help to accelerate progress in improving maternal and child nutrition? *The Lancet*, 382(9891):536 – 551.

Swindale, A. and Bilinsky, P. (2006). Development of a universally applicable household food insecurity measurement tool: process, current status, and outstanding issues. *The Journal of nutrition*, 136(5):1449S–1452S.

Taye, H. (2013). Evaluating the impact of agricultural extension programmes in sub-Saharan Africa: Challenges and prospects. *African Evaluation Journal*, 1(1).

Udry, C. (2010). The economics of agriculture in Africa: Notes toward a research program. *African Journal of Agricultural and Resource Economics*, 5(1).

Van der Ban, A. and Hawkins, H. (1996). *Agricultural extension*. Blackwell Science, Oxford Malden, MA.

Vyas, S. and Kumaranayake, L. (2006). Constructing socio-economic status indices: how to use principal components analysis. *Health policy and planning*, 21(6):459–468.

World Bank, T. (2007). *World development report 2008 agriculture and development*. World Bank Eurospan distributor, Washington, D.C. London.

CHAPTER 4

Lottery - and survey-based risk attitudes linked through a multichoice elicitation task [1]

[1]This chapter is based on a joint work with Giuseppe Attanasi (BETA, University of Strasbourg), Nikolaos Georgantzís (University of Reading) and Daria Vigani (Catholic University of Milan). It has been written during a visiting period at the University of Strasbourg. We acknowledge the project "Attractivité IDEX 2013" for financial support.

4.1 Introduction

Experimental evidence and, thus, recent theories of individual decision making have acknowledged the fact there is more to a decision maker's attitude to risk than risk aversion alone. Concepts like loss aversion, reference points or aspiration levels, probability weighting and even violations of stochastic dominance are non-negligible aspects which are jointly or separately accounted for by modern theories aiming at accommodating previously disturbing and paradoxical phenomena.

However, several practitioners and the vast majority of experimental researchers seem to rely on risk aversion alone, when (the former) pricing a risky asset and (the latter) eliciting subjects' risk attitudes as an explanatory variable of their behavior in another decision making context. For example, in more than half of the occasions in which experimental economists wish to account for their subjects' risk attitudes as a primary or secondary aspect of their behavior, the Holt and Laury (2002) – *HL* hereafter – procedure is adopted, which is primarily a uni-dimensional test often used to map decisions on a uni-parametric utility function. A different procedure involves a survey question asking subjects to assess their attitude towards risk (self-assessed risk attitude). Interestingly, whether this is done by a single question or with a more complex test like Zuckerman's Sensation Seeking Scale (Zuckerman, 1994), even in this literature the variable used is a uni-dimensional construct assessing a person's overall riskiness.

In this paper, we contrast both aforementioned methods, to choices in a Lottery-Panel Test (Sabater-Grande and Georgantzis, 2002) – *SG* hereafter –, this method disclosing more information on a decision maker's risk attitudes. Specifically, each participant makes a choice among a series (panel) of alternative lotteries. Four panels are constructed, each one of which provides subjects with a different incentive (risk premium) to make riskier choices. A parametric approach to the test can offer a simple prediction on subjects' behavior across panels and is easily comparable with uni-dimensional mapping on the utility parameter space like in *HL*. The richness of patterns emerging as deviations from the expected-utility predicted behavior across panels, al-

lows us to classify subjects according to criteria which are not applicable in simple models.

A rather surprising finding that is recurrently reported by different experimental studies is that risk attitudes elicited through different methods differ significantly from each other.[2] Several aspects of this finding relate to the mere nature of the tasks. For example, tasks with losses are naturally expected to capture different dimensions of subjects' psychological attitudes as compared to tasks limited to the gains domain. Furthermore, it is rather easily accepted that tasks covering different payoff ranges would also lead to significant differences in the elicited attitudes.

The consequences of accepting such differences as natural can be of two types: 1. those concerning the relation between the elicitation task and the underlying theoretical decision-making model under risk, and 2. those related to the usefulness of the task as a method of obtaining an explanatory variable to empirically capture the role of subjects' risk attitude on his/her behavior in a different task. Both issues are largely neglected, not so much by the studies specifically designed to compare risk attitudes elicited through different tasks, but by those acting as *simple users* of the tasks as a method of generating a risk-attitude related explanatory variable for their primary data from an experiment.

Regarding the first issue, the most striking feature of a number of broadly used tasks is their dependence on a single choice made by each subject. It is straightforward to see why such a strategy is both tautologically consistent with any uniparametric description of the decision problem solved by the decision maker, and in dissonance with all modern theories based by definition on more than the product of probabilities with the uniparametric utility transformation of the associated monetary outcomes.[3]

Regarding the second issue, we feel that it can be, at the same time, more

[2]For a recent example of five elicitation methods and reference to such results, see Crosetto and Filippin (2015). As in previous experimental studies on the topic, they also find that the estimated risk aversion parameters vary greatly across tasks.

[3]Anecdotally, we would like to refer to the case of a referee stating and an editor agreeing that "to elicit one's risk parameter from a single choice is not problematic, whereas to obtain it from many decisions generally leads to inconsistencies".

urgent to address and less problematic. This is so, because there seems to be some consensus on the intuitive but not sufficiently supported fact that decisions made across similar tasks should be expected to elicit attitudes which do not significantly differ from each other. Of course, one should not forget that even the repetition of exactly the same task by the same individual would most probably lead to differences. But such differences follow specific patterns, some of which have been documented empirically,[4] and even conform to the well-known paradigm of preference imprecision.[5]

Thus, even when choices by the same subject in the same task over different trials are different, attitudes elicited in similar tasks should be related to some extent, even through correlation of the ranking that subjects received by their choices in the overall population. If this *desideratum* is satisfied, then eliciting risk attitudes as an explanatory variable of behavior in another task can be considered a meaningful strategy. On the contrary, if any arbitrarily small change of the context produces different attitude elicitations that are not systematically related across tasks, we risk failing to satisfactorily answer to the question "*does any of what we are observing in the lab relate at all with what anyone (even the same subject) does outside the lab?*"

In this paper, we aim at shedding light on the reliability of *HL*, the method mainly used in the last decade to elicit risk attitudes across a wide array of contexts and environments. To achieve this goal, we compare it experimentally to another risk-elicitation method, *SG*, which is made by a series of four tasks that we think can help us in identifying risk-related attitudes disregarded by the implementation of *HL* alone. Using the type classification

[4]For example, *regression to the mean* has been found to affect repeated choices in the same task by García-Gallego et al. (2011). Furthermore, Lévy-Garboua et al. (2012) found a significantly higher elicited risk aversion in sequential than in simultaneous treatment, in decreasing and random than in increasing treatment, in high than in low-payoff condition. Their findings suggest that subjects use available information that has no value for normative theories. Cox et al. (2014) have rationalized some of these findings by showing the role of the payment mechanism in these distortions. Indeed, they find that random-lottery incentive mechanisms – as those usually employed in risk-elicitation tasks – may decrease the proportion of risky choices in the population, if compared to a one-task design. This could explain why significantly more risk aversion emerges under multiple-task than under one-task elicitation methods.

[5]See Butler and Loomes (2007).

emerging from the choices made in *SG*, we want to see 1. whether the correlation between the risk-aversion orderings under the two elicitation procedures increases, and 2. whether any of these two monetary-incentivized mechanism is a good predictor of self-assessed risk attitudes (e.g. elicited through a hypothetical question about one's general willingness to take risks).

We report two rather exceptionally positive findings which can contribute to a literature full of negative or contrasting results. First, we find evidence of some external validity of these two mutually uncorrelated risk-attitude elicitation methods – *HL* and *SG* – as predictors of self-reported risk attitudes in general human domains. Second, and more importantly, we show that sufficiently similar incentivized mechanisms elicit correlated decisions in terms of monetary risk aversion *only if* other risk-related attitudes are disentangled. Considered together, our results indicate that, whereas both *HL* and *SG* are reasonably good predictors of self-assessed risk attitudes, the use of a more complete description of subjects' risk attitudes is helpful when stating the ability of each test to predict self-reported attitudes.

The rest of the paper is structured as follows. In Section 4.2 we review the literature on risk-aversion elicitation, and show more in depth the specific issues on which our study aims at contributing. Section 4.3 presents our experimental design. In Section 4.4 specific behavioral hypotheses are introduced. Section 4.5 analyzes the experimental results, which are discussed in the concluding section.

4.2 Literature review

Assessing and measuring individuals' risk preferences is a fundamental issue for economic analysis and policy prescriptions (Charness et al., 2013). As a result, economists and other social scientists have developed a wide variety of experimental methodologies to elicit individual risk attitudes.

Risk preferences have been indirectly derived from *first-price sealed bid auctions* (e.g. Cox et al., 1982, 1985, 1988), or elicited as *lottery certainty equivalents* (Becker et al., 1964). Individual degrees of risk aversion have

also been experimentally measured through asking subjects to input a value for one of the outcomes of a lottery that would make them indifferent with respect to another proposed lottery (the so-called *trade-off method* by Wakker and Deneffe, 1996).

Survey methods have also been employed, where subjects are asked to *self-report* their risk preferences through a series of *hypothetical questions* concerning a general willingness to take risks (see, e.g., Dohmen et al. (2011), for a representative sample of roughly 22,000 German subjects) or a specific willingness to participate in a lottery (see, e.g., Attanasi et al. (2013), for a sample of about 10,000 Italian subjects over five consecutive years).

Nowadays, among economists, the most common and widespread procedure to measure risk preferences in the laboratory is to ask subjects to choose one lottery (*single decision*) among a panel of lotteries. These lotteries can either entail a single choice among a set of predetermined prospects, presented in an abstract way (Eckel and Grossman, 2008), or can be framed as an investment decision (Charness and Gneezy, 2010), or still can be presented by means of a visual task, without making any explicit reference to probabilities (Lejuez et al., 2002; Crosetto and Filippin, 2013).

As an extension of the previous method, subjects were asked to make *multiple decisions* between pairs or panels of risky lotteries. This is the case under investigation in this paper. As a matter of fact, both Holt and Laury (2002) and Sabater-Grande and Georgantzis (2002) use this last method in order to elicit risk aversion. Holt and Laury (2002) – and follow-up papers[6] – is the most well-known example of a "multiple price list design" which, according to Cox and Harrison (2008), was first used in Miller et al. (1969). The risk-elicitation procedure used in Sabater-Grande and Georgantzis (2002) is a slightly revised version of the "ternary lotteries approach" (see, e.g., Roth and Malouf, 1979).

In this paper the two risk-elicitation methods are proposed in a within-subject design. Besides the original version, the multiple pairwise compar-

[6]See Harrison et al. (2005) and Holt and Laury (2005): the former demonstrated and the latter confirmed the possibility of order effects in Holt and Laury (2002) original design by scaling up real payments by 10 or 20 times.

ison in *HL* has been usually implemented with two non-mutually-exclusive variants. The first one – "switching multiple price list" – was introduced by Harrison et al. (2005) and studied at length by Andersen et al. (2006): Monotonicity is enforced, i.e. the subject is asked to pick the switch point from one lottery to the other and non-switch choices are filled in automatically. The second one concerns doubling the number of outcome probabilities for which the two lotteries are compared, in order to allow the subject to make choices from refined options: This second variant – together with enforced monotonicity – has been implemented by Attanasi et al. (2014a), where *HL* is made of 20 lottery pairs instead of 10. The second risk-elicitation method is exactly the one used in Sabater-Grande and Georgantzis (2002) and in follow-up papers.[7]

As anticipated in the Introduction, our choice of analyzing and comparing two alternative methods of measuring risk preferences is partly due to a puzzling result in experimental economics: The degree of risk aversion shown by subjects in the laboratory is often varying across different elicitation techniques (see, e.g., Isaac and James, 2000; Dave et al., 2010), although some correlations are found among monetary-incentivized instruments and survey-based methods (see, e.g., Vieider et al., 2015).

In recent years, a growing literature is investigating different risk-elicitation methods, comparing their effectiveness in eliciting risk attitudes in non-interactive settings. Harrison and Rutström (2008) review experimental evidence on risk aversion in controlled laboratory experiments. The authors examine the experimental design of several procedures that allow direct estimation of risk preferences from subjects' choices, as well as the way to draw inference about laboratory behavior. Furthermore, they provide an investigation on how the data generated by these procedures should be analyzed. In the same line, Charness et al. (2013) provide a discussion of a series of prevailing methods for eliciting risk preferences. They outline the strengths and weaknesses of each of these methods. In particular, they highlight that choosing which method to utilize is largely dependent on the question the researcher wants to answer. Both these reviews of risk-elicitation methods

[7]See Georgantzís and Navarro-Martínez (2010), and García-Gallego et al. (2011).

include a thorough discussion of *HL*.

Among experimental studies that compare *HL* with other risk-elicitation methods, two are relevant for our paper. Charness and Viceisza (2015) compare two incentivized risk-elicitation methods in a between-subject design, namely *HL*, and the modified version of the Gneezy-Potters method as presented in Charness and Gneezy (2010). In both treatments, subjects also self-report their risk attitude by answering a hypothetical question similar to the one in Dohmen et al. (2011). The experiment was run in rural Senegal, with the aim of providing guidance to experimenters wishing to use risk-elicitation mechanisms in the rural developing world. Crosetto and Filippin (2015) compare five incentivized risk-elicitation methods in a between-subject design: *HL*, Eckel and Grossman (2008), Charness and Gneezy (2010),Lejuez et al. (2002), and Crosetto and Filippin (2013). All experimental sessions being run in Jena (Germany), they find that subjects' estimated risk aversion parameters vary greatly across tasks.

Our work is positioned exactly in this branch of the literature.

Using original data from a homogeneous population of Italian subjects, we provide an experimental comparison of *HL* with another incentivized risk-elicitation method (*SG*), and a self-assessment measure of risk attitude (a hypothetical question similar to the one in Dohmen et al., 2011). Differently from the previous literature, this comparison is made through a *within-subject* design: each subject in our experiment goes through the three risk-elicitation procedures, and we control for other effects. Furthermore, we compare two *multiple-decision* methods, while in previous studies *HL* has only been compared to single-decision methods. In fact, we think that coupling *HL* with another multiple-decision mechanism could help shed more light on the reliability of the former.

As underlined above, *HL* is the most widely used risk-elicitation method in experimental economic analyses in the last ten years: When risk aversion is considered as an explanatory variable for subject's behavior in an individual or strategic decision setting, a preliminary test of risk aversion (preliminary with respect to the main decision setting where subjects' behavior should be analyzed) is needed. In this regard, *HL* has a clear advantage with respect to

many other risk-elicitation methods: Especially when enforcing monotonicity – as it is more frequently the case in economic experiments – it allows to completely describe a subject's risk attitude through just one subject's choice.

It is well known that this requires assuming that the subject is a von Neumann-Morgenstern expected utility maximizer. Many papers have shown that this assumption is questionable (see Wakker (2010) for a review), although other models (e.g., prospect theory) do not seem to have significantly higher explanatory power than expected utility (Harrison and Rutström, 2009). However, the goal of our exercise is not to test the expected utility assumption in *HL*. Rather, we are interested in other risk-related attitudes that are not taken into account when analyzing *HL* data, since with just one choice per subject, by construction, attention is restricted to the curvature of the uniparametric (Bernoullian) utility function.[8] Therefore, other relevant risk-related attitudes may be disregarded.

The intuition behind our exercise is that the series of four tasks that constitute *SG* can help us in identifying some of these further risk-related attitudes. This is the main reason why we focus on the comparison between *HL* and *SG*. Using the type classification emerging from the choices made in *SG*, we want to see whether the correlation between the risk-aversion orderings under the two elicitation procedures increases. Furthermore, and more importantly, we want to check whether, by disentangling subjects according to these further attitudes, the correlation between a subject's self-reported sensitivity to risk and the monetary-incentivized choice made respectively in *HL* and in *SG* is higher. A positive answer to the last question – that is what we actually found in this paper – should help explain why experimental economists (rather than psychologists) usually do not rely on self-reported measures of risk: The hypothetical questions used to let subjects self-assess their level of risk attitude might hide risk-related motivations other than

[8]Notice that this problem would emerge also in the absence of enforced monotonicity. In fact, when *HL* is performed – as in the original paper – by asking subjects to make a choice between the two options for each of the 10 outcome probabilities, a subject who switches from one lottery to the other more than once as the probability of the best outcome increases, is still considered as if he/she has made just one choice (one switch). This is done by assigning to this subject as switch point from one lottery to the other the one corresponding to the number of safe choices the subject has made.

monetary risk aversion, i.e., the curvature of the uniparametric expected utility function.

4.3 Experimental design

Participants were undergraduate students in Economics, recruited at Bocconi University in Milan on October 2013. The number of participants was 62. Each subject could only participate in one session: Two sessions were run with 31 subjects each in a computerized room of Bocconi University, with subjects being seated at spaced intervals.

In each session, subjects faced two risk-elicitation tasks (*HL* and *SG*) on a within-subject base. In the two sessions, the two tasks were shown in reverse order. Only one of the two tasks was used to determine subjects' final earnings: The choice of the task to be paid was made in a random way, by flipping a coin. Payment was preceded by a questionnaire, which included a question about self-assessment of risk attitude.

Average earnings were €15.90, including a €3.00 show-up fee. The average duration of a session was 45 minutes, including instructions and payment. The experiment was programmed and implemented using the z-Tree software (Fischbacher, 2007).

4.3.1 Task 1 (*HL*)

The first task was *HL* with the two variants of enforced monotonicity and 20 (rather than 10) lottery pairs, and set of lottery payoffs as in Attanasi et al. (2014a). See Figure A1 in the Appendix.

Features of the task

- Subjects were presented with a battery of 19 pairs of two-outcome lotteries, numbered from line $L1$ to line $L19$, and a last (empty) line $L20$ (bottom line of Figure A1).

- Each pair described two lotteries called A and B.

- Each lottery presented two positive monetary outcomes and their associated probabilities.

- The two monetary outcomes of each lottery were kept constant: For each line $L1$–$L19$, lottery A always had the two outcomes , $\bar{x}_A =$€ 12.00, $\underline{x}_A =$€ 10.00, and lottery B always had the two outcomes, $\bar{x}_B =$€ 22.00, $\underline{x}_B =$€ 0.50 .

- Within each pair, $\bar{x}_A$ and $\bar{x}_B$ were attached the same probability p, with p increasing – gradually and monotonically – when moving from the top ($L1$) to the bottom ($L19$) of the battery of lottery pairs.

- Probabilities were framed by means of an urn that contained 20 tickets, numbered from 1 to 20, the number of tickets associated to the highest of the two outcomes, $\bar{x}_k$, being independent of the lottery ($k = A, B$) and varying with the line. In particular, in $L1$ the highest outcome was assigned ticket no. 1; in $L2$, tickets no. 1 and 2; ... ; in $L19$, all tickets but no. 20. Hence, in the light of a final random draw of a ticket from the urn, the probabilities of $\bar{x}_k$ and of $\underline{x}_k$ were respectively: 1/20 and 19/20 in $L1$; 2/20 and 18/20 in $L2$; ... ; 19/20 and 1/20 in $L19$.

What subjects were asked to do

Given the battery of lotteries, each subject was asked to choose the *switch line*, i.e. the pair of lotteries starting from which he/she preferred lottery B to lottery A. Thus, for all pairs of lotteries above the switch line, a subject preferred lottery A to lottery B, while starting from the pair on the switch line and for all the pairs below, he/she preferred lottery B to lottery A. A subject preferring lottery A to lottery B for all the 19 pairs, selected the last (empty) line $L20$.

Determination of the subject's earnings

Suppose that task 1 was randomly selected (by flipping a coin) at the end of the experiment to determine subjects' earnings. Then, for each subject the computer would randomly select a pair of lotteries, i.e. one of the 19 lines of the battery of lotteries. The *randomly-selected line* indicated the number of tickets assigned to the highest outcome, hence the probability associated to the two outcomes of both lottery A and lottery B.

If a subject's switch line was below the randomly-selected line, then the two lottery outcomes for which that subject played were $\bar{x}_A =$€ 12.00 and $\underline{x}_A =$€ 10.00; otherwise, the two lottery outcomes for which that subject played were $\bar{x}_B =$€ 22.00 and $\underline{x}_B =$€ 0.50.

Then, an experimenter randomly drew one of the 20 tickets contained in the urn (physical implementation).[9] The ticket drawn by the experimenter was used to determine whether each subject earned the higher or the lower outcome of the chosen lottery in the randomly-selected line.

4.3.2 Task 2 (SG)

The second task was SG as implemented by – among other studies – Georgantzís and Navarro-Martínez (2010), García Gallego et al. (2012) and García-Gallego et al. (2011).[10] See Figure A2 in the Appendix.

Features of the task

- Subjects faced four decision problems. Each problem concerned a panel of 10 two-outcome lotteries described in three rows. Each lottery had a positive outcome X and a null outcome.

[9] See Section 5.3 of Attanasi et al. (2014b) for the pros and cons of physical vs. computerized instruments when generating realizations of random processes in laboratory experiments on decision under uncertainty.

[10] In Sabater-Grande and Georgantzis (2002) the payoffs are expressed in pesetas, since the experimental sessions were run in Spain before the introduction of the Euro as official currency in the European Union. In the follow-up studies cited above, still run in Spain, the payoffs are equivalently expressed in Euros.

- The first row presented, for each of the 10 lotteries, the probability p assigned to the positive outcome X.

- The second row presented, for each of the 10 lotteries, the positive outcome X.

- The third row consisted of 10 empty cells, for each subject to indicate with a cross the preferred lottery in each of the four panels.

- Across each panel of lotteries, neither X nor p were kept constant. However, p was the same for the same column of each panel: Probabilities were framed by means of an urn that contained 10 tickets, numbered from 1 to 10, the number of tickets associated to X decreasing with the column number of each panel of lotteries. In particular, in the first column (leftmost lottery), all tickets were associated to X; in the second column, all tickets but no. 10 were associated to X; ...; in the tenth column (rightmost lottery), only ticket no. 1 was associated to X. Hence, in the light of a final random draw of a ticket from the urn, for each panel, the leftmost lottery represented the safest option ($p = 100\%$) with the lowest positive outcome, while the rightmost lottery represented the riskiest option ($p = 10\%$) with the highest positive outcome. Moving from the left side to the right side in a panel, the lotteries were constructed in order to compensate riskier options with increases in the expected payoff pX. Thus, the positive outcome X increases with the column number of each panel of lotteries.

- Formally, each continuum of lotteries was defined by the pair (c, t) corresponding, respectively, to the certain payoff c above which the expected payoff of the lottery L was increased by $t > 0$, times the probability of earning nothing, i.e.

$$E(L) = pX = c + (1 - p)t \quad \Rightarrow \quad X(p) = \frac{c + (1 - p)t}{p} \ ,$$

where t is a panel-specific risk premium, which generates an increase in the lotteries' expected values as one moves from safer option (left

side of Figure A2) to riskier (right side of Figure A2) options within the same panel.

- In particular, the four panels of lotteries were constructed using $c = €1$ in all panels and $t = 0.1$ for panel 1, $t = 1$ for panel 2, $t = 5$ for panel 3, and $t = 10$ for panel 4. Hence, the above formula shows that in the first column (leftmost lottery) the certain positive outcome was the same for each panel ($X(100\%) = c = €1$), while in all remaining columns the certain positive outcome was increasing in the panel number, being maximum in the last column (rightmost lottery), with $X(10\%)$ being €10.90 for panel 1, €19.00 for panel 2, €50 for panel 3, and €100 for panel 4.

What subjects were asked to do

For each panel of lotteries, each subject was asked to choose one of the 10 lotteries (X, p) that implied a probability p of earning X, else nothing. Hence, for each of the four panels, a subject was asked to put a cross in the empty cell corresponding to his/her preferred lottery among the 10 available lotteries (10 columns).

Determination of the subject's earnings

Suppose that task 2 was randomly selected (by flipping a coin) at the end of the experiment to determine subjects' earnings. Then, for each subject the computer would randomly select one of the four panels of lotteries.

For the randomly-selected panel of lotteries, the cross put by a subject in the empty cell corresponding to his/her preferred lottery indicated the positive outcome of the lottery (X) and the number of tickets assigned to this outcome (p).

Then, an experimenter randomly drew one of the 10 tickets contained in the urn (physical implementation). The ticket drawn by the experimenter was used to determine whether each subject earned the positive or the null outcome of his/her preferred lottery in the randomly-selected panel.

4.3.3 Questionnaire

A questionnaire about some idiosyncratic features has been submitted at the end of the experiment. Each subject was asked his/her gender, age, year and field of study, previous attendance of an advanced course in Decision/Game Theory, and a question about self-assessment of general attitude towards risk, similar to the one used in Dohmen et al. (2011). In particular, the question was posed using the same wording of Bernasconi et al. (2014), which also run their experiments with Italian subjects: "In a scale from 1 to 10, how would you rate your attitude towards risk: are you a person always avoiding risk or do you love risk-taking behavior?", where 1 was associated with the statement "I always choose the safest option and try to avoid any possible risk" and 10 referred to "I love risk and I always choose the more risky alternative".

4.4 Behavioral Hypotheses

Each task is mainly targeted to elicit a subject's degree of (monetary) risk aversion, through a different method. However, task 1 (*HL*) being characterized by less "flexibility" in the subject's available choices with respect to task 2 (*SG*), the latter can be used to disclose and disentangle other risk-related motivations.

In fact, while in *HL* a subject is repeatedly asked to choose the preferred pair of lottery-outcomes between two pairs with *the same* associated probabilities, in each panel of lotteries in *SG* the subject is asked to pick the preferred outcome-probability combination, with both the positive outcome and its associated probability being *different* for each lottery. With this in mind, our **first aim** is to use a subject's four choices in *SG* to disentangle his/her risk-related motivations behind the unique choice (switch line) in *HL*.

A Constant Relative Risk Averse (*CRRA* hereafter) utility function of the form $U(x) = \frac{x^{1-r}}{1-r}$ is assumed to elicit a subject's (monetary) risk attitude in both *HL* and *SG*, implying risk aversion for $r > 0$, risk neutrality for $r = 0$, and risk proneness for $r < 0$.

In *HL*, given the structure of the battery of lotteries (see Figure A1 in the Appendix), the higher the number of the switch line (pairwise comparison at which a subject chooses to switch from lottery A to lottery B), the higher his/her disclosed degree r of relative risk aversion (see Table A1 in the Appendix). In particular: a switch line from $L1$ to $L9$ would reveal risk proneness (the smaller the number of the switch line, the higher $|r|$, the degree of risk proneness); a risk-neutral subject would indicate $L10$ as switch line; a switch line from $L11$ to $L19$ would reveal risk aversion (the greater the number of the switch line, the higher r, the degree of risk aversion).

In the original version of *HL*, subjects had to choose the preferred lottery between A and B in each of the 10 lines of the battery, giving rise to the possibility of inconsistent behavior at the individual level.[11] In our study, due to the *enforced monotonicity* feature of the implemented variant of *HL*, consistency has been imposed: Picking the switching line directly provided an interval estimate of the subjects' coefficient r of relative risk aversion. Moreover, *doubling the number of outcome probabilities* for which lotteries A and B are compared (20 lottery pairs instead of 10) allowed a more precise interval estimate of r, given the switching line.

As far as *SG* is concerned (see Figure A2 in the Appendix), an expected-utility maximizing subject with a *CRRA* utility function as introduced above would choose a lottery (X^*, p^*) with $p^* = \frac{cr}{t} + r$ in each of the four panel of lotteries. Hence, the chosen probability p^* of the positive outcome is monotonically decreasing in the subject's degree r of relative risk aversion: Safer choices in each panel (left side of each panel in Figure A2) are associated with a higher r (see Table A2 in the Appendix). In particular, all risk-neutral and risk-loving subjects should choose the lottery at the far right extreme of each panel in Figure A2 ($p = 0.1$ in Table A2). Furthermore, given that the panel-specific risk premium t increases by construction with the panel index, all *CRRA* subjects with a given r should not choose safer lotteries (weakly monotonic transitions) as they move from panel-1 lotteries to panel-

[11] As a matter of fact, individuals going back and forth in their choices could be considered inconsistent with a *CRRA* pattern: This happened for around 13% (7%) of subjects in the initial (final) low-payoff task of *HL*.

4 lotteries. In terms of Table A2, moving to a panel with a higher index, *CRRA* subjects should choose in this panel a lottery not being on the left side of the lottery chosen in the previous panel.

SG has several advantages that are useful to our analysis. Firstly, the above-mentioned theoretical predictions also hold for other well-known utility functions like *CARA* (Constant Absolute Risk Aversion) or other functional forms for *CRRA*, different from the one for which the elicited r in Table A1 (for *HL*) and Table A2 (for *SG*) has been calculated. Secondly, *SG* exposes a subject to the same wide range of probabilities in each panel-wise comparison, and to a systematic spectrum of monetary rewards from €1 (far left extreme of each panel) to the relatively high payoff of €100 (far right extreme of the panel 4). Finally, the test offers a range of different returns to risk so that a highly-risk-averse subject might refuse to take too risky options when a higher return is at stake (e.g., he/she chooses $p = 0.4$ in panels 3 and 4), while he/she could be attracted by highly-risky prospects when returns are lower (e.g., he/she chooses $p = 0.1$ in panels 1 and 2). This is incoherent with the *CRRA* assumption, however it can disclose other interesting risk-related motivations, as we will see in the next section. Thus, unlike all uni-dimensional tests of monetary risk attitude, *SG* may be used to classify subjects not only according to their willingness to take monetary risks, but also with respect to their propensity to change their "objective function" across different risk-return combinations. This would help disentangle risk-related motivations that might explain a subject's different choice in *HL*.

Our **second aim** is to check whether the self-reported assessment of risk preferences in the final questionnaire is an explanatory variable for the degree of risk aversion elicited in any of the two monetary-incentivized tasks of our experiment. Many field studies have shown that asking a general hypothetical question about the self-assessment of risk aversion is a simple procedure to estimate risk attitudes of subjects (see, among others, Guiso and Paiella, 2008). In particular, Dohmen et al. (2011) have shown that such questions are as effective as other common and much more complicated procedures used in laboratory experiments.

In this paper, we separately check whether the answers to the hypothetical

question on the self-assessment of the degree of risk aversion relate with the subject's choices in *HL* and in *SG*. Notice that in the general hypothetical question proposed in our questionnaire, the higher the selected number in the 1-10 scale, the lower the subject's *self-assessed degree of risk-aversion*. Hence, this should correlate negatively with the number of the switching line in *HL* and positively with the probability chosen in each panel of *SG*.

Given the relevance of *HL* for current laboratory experiments on risk elicitation, we further focus on it: We use a subject's answers to the general hypothetical question on risk assessment to disentangle his/her risk-related motivations behind the choice in *HL*. Therefore, in the final part of the next section, both *SG* and the hypothetical question – together with the other questions on idiosyncratic features in the final questionnaire – will be used as regressors in the analysis of behavior in *HL*.

4.5 Results

4.5.1 Aggregate analysis

We observe no significant effect of proposing *HL* before *SG* or showing them in reverse order. Thus, in the following we will pool data from the two treatments.

The distribution of individuals among their risk-related choices is quite close in the two tasks: 74% of subjects disclose risk aversion in *HL*, and 77% in *SG* (on average over the four panels) disclose risk aversion. However, this first check is made only at a between-subject level. We must also check whether, within-subject in the two tasks, the sign of the risk attitude does not change, i.e. if a subject showing risk aversion (proneness) in *HL* also shows risk aversion (proneness) in each of the four panels of *SG*.

Table 4.1 reports the conditions to be satisfied to pass such test, by summarizing the information reported in Tables A1–A2 in the Appendix (we indicate with SG_i the choice made in panel i of *SG*, with panel number $i =$ 1,2,3,4). Indeed, a subject disclosing risk aversion because switching after

$L10$ in HL (46/62, 74% of the sample), should choose a lottery with $p \geq 0.4$ in panel 1, and a lottery with $p \geq 0.2$ in the other three panels of SG. All other subjects should choose a lottery with $p \leq 0.3$ in panel 1, and a lottery with $p = 0.1$ in the other three panels of SG.

Table 4.1: Threshold levels for coherent choices according to the sign of r

	Switch Line	*Chosen Probability*	
	HL	$\boldsymbol{SG_1}$	$\boldsymbol{SG_2}$-$\boldsymbol{SG_4}$
Risk averse ($r > 0.038$)	$L11$-$L20$	$0.4 \leq p \leq 1$	$0.2 \leq p \leq 1$
Risk loving and risk neutral ($r < 0.038$)	$L1$-$L10$	$p \leq 0.3$	$p = 0.1$

Figure 4.1 reports the percentage of subjects disclosing risk aversion according to the elicited r in HL, which also show a positive r in each of the four panels of SG, and in all the four panels considered together. We can see that, apart from panel 1 – where the payoff scale is much smaller than in HL[12] – the majority of HL-risk-averse subjects (always more than 60% in each of the three last panels) also show a risk-averse behavior in SG. Conversely, among the few (16/62, 26% of the sample) HL-risk-neutral and HL-risk-loving subjects, almost none discloses the same sign of risk attitude in any panel of SG. For example, in panels 2 and 3, they all make risk-averse choices. All these findings are summarized in Result 1.

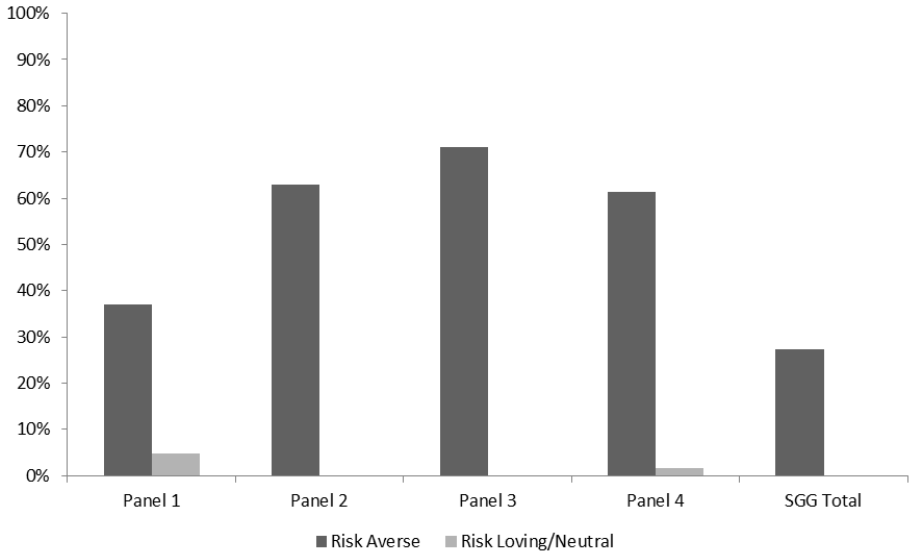
Result 1. *The majority of subjects showing risk aversion in HL also show risk aversion in the four panels of SG. Risk-neutral and risk-loving subjects according to HL disclose risk aversion in the four panels of SG.*

We now check whether the ordering of subjects' risk preferences does not vary too much from one task to another.

The controls for SG work as they should: A significant positive correlation (at least 50%, always significant at the 1% level) is found among the ordering of any randomly-chosen pair of panels of SG. Furthermore, the level of accepted risk decreases (on average across all subjects) with the panel number, coherently with the assumption of $CRRA$ (moving from panel 1 to

[12]Indeed, 50% of all subjects (31/62) choose a lottery with $p \geq 0.4$ in panel 1, thereby disclosing risk neutrality or proneness.

Figure 4.1: Subjects showing the same sign of risk attitude in *HL* and *SG*



panel 4 we have increasing stakes for the same number of tickets assigned to the positive outcome).

We perform the Spearman rank correlation test among choices made in each of the four panels in *SG*, and choices in *HL*, and between the latter and a variable representing the average choice in the four panels (SG_{Avg}). Results are reported in Table 4.2.

Recall that: In *HL*, the larger the number of the switch line, the *smaller* the number of tickets assigned at this line to the lower of the two outcomes (see Figure A1 in the Appendix), and the *higher* the subject's disclosed degree of risk aversion (see Table A1 in the Appendix); in each panel of *SG*, more in the left the chosen lottery is, the *smaller* the number of tickets assigned to the null outcome (see Figure A2 in the Appendix), and the *higher* the subject's disclosed degree of risk aversion (see Table A2 in the Appendix). Therefore, from now on, when analyzing results for *HL*, we consider as index the *number of tickets assigned in the switch line to the lower outcome*; when analyzing results for each panel of *SG*, we consider as index the *number of tickets assigned in the chosen lottery to the null outcome*. Hence, a positive correlation between these two indexes would mirror the positive correlation

between disclosed risk attitudes in the two tasks.

Table 4.2: Rank correlations between self-assessed risk and average choices in the two tasks, by panel.

	$HL - SG_1$	$HL - SG_2$	$HL - SG_3$	$HL - SG_4$	$HL - SG_{Avg}$
Spearman's rho	0.11	0.04	0.11	0.04	0.13
p-value	0.39	0.78	0.37	0.76	0.31

Rank correlations in Table 4.2 lead to the following:

Result 2. *A positive but small and not significant correlation is found between subjects' risk ordering in HL and their risk ordering in any of the four panels of SG.*

Thus, different risk-elicitation instruments seem to lead to different orderings of the relative risk aversion coefficient r, if Expected Utility is assumed for all subjects. Note that this result still holds when conditioning on age, gender, and for past attendance of a course in decision/game theory. However, a positive and quite surprising finding emerges if looking at subjects' self-reported risk through the hypothetical question in the final questionnaire.

In particular, we make use of the self-assessed risk variable in order to check on the rank correlation between this subjective measure and the risk-related choices made by the subjects in the two tasks. Recall that in the question about self-assessment of general risk attitude, the *smaller* (closer to 1) the chosen number, the *higher* the self-assessed general aversion to risk. Hence, a positive correlation between this choice and the above defined index in a risk-elicitation task (*HL* or *SG*) would mirror the positive correlation between self-assessment of risk and the disclosed monetary risk attitude in the task.
This check leads to the following:

Result 3. *Significant positive correlation is found between subjects' ordering expressed by self-reported risk and the risk ordering in HL (rho = 0.47, p-value = 0.000). Significant positive correlation is also found between*

the ordering expressed by self-reported risk and the risk ordering in SG (rho = 0.48, p-value = 0.000).

As can be noticed, the two correlation coefficients are very close. Both methods seem to be able to account for a good amount of inter-individual differences in general aversion to risk, with similarly high explanatory power.

From these preliminary results, three questions arise:

1) *Why the rankings produced by the two methods are not correlated while instead each of them is correlated with self-assessed general aversion to risk?*

2) *Why the correlation coefficient of each method with self-assessed general aversion to risk is smaller than 50%?*

3) *Why are the correlation coefficients of each method with self-assessed general aversion to risk so close?*

The analysis in the next subsection, which account for both idiosyncratic features (elicited in the questionnaire) and other risk-related motivations (as emerging from choices in the four panels of *SG*) is meant to answer the above questions. The reliability of *HL* as instrument for risk-aversion elicitation crucially depends on the answers to the previous questions.

4.5.2 Type classification analysis

Results 1–3 above lead us to think that there can be other subjects' features and motivations (other than monetary risk aversion) orienting subjects' choices in each of the two analyzed instruments. If we disentangle subjects according to these motivations, we should find individuals who better disclose their self-reported risk aversion in *HL* and others who better disclose it in *SG*.

First, we check whether idiosyncratic features (gender, age, education, etc.), elicited through the final questionnaire, are of some help in providing a coherent explanation for the previous findings.

Furthermore, since *HL* only requires one choice for each subject, while *SG* requires four choices for each subject, we use this second instrument in order to disentangle risk-related motivations.

Individual characteristics

We run again the Spearman rank correlation tests between self-assessed risk and individuals' choice in both *HL* and the average choices in *SG* by sub-groups of population.

First of all, we can look at **gender**. We find that females show a significantly higher correlation between self-assessed risk and risk behavior in both *HL* and SG_{Avg}. The two correlation coefficients are again of the same magnitude: $rho = 0.60$ for *HL* (p-*value* $= 0.038$), $rho = 0.60$ for *SG* (p-*value* $= 0.035$). The correlation coefficients for the sub-group of males, though significant, are lower than those obtained for the whole sample: $rho = 0.31$ for *HL* (p-*value* $= 0.050$), $rho = 0.42$ for *SG* (p-*value* $= 0.006$).

Another interesting issue is whether having attended an advanced course in decision or game theory could strengthen the correspondence between subjects' self-assessed risk and their actual risk-related choices in the two tasks. Our auxiliary assumption is that such attendance should indicate some **background in mathematically-related disciplines** (recall that our subjects are undergraduate students in Economics). The usual test reveals that the previous attendance of a decision/game theory course increases the correlation between self-reported risk and average choice in *SG*, while the opposite effect is found with regards to *HL* (see Table 4.3).

Table 4.3: Rank correlations between self-assessed risk and the two tasks, disentangled by gender and backgroungd in decision/game theory

	Female	***Male***
Self-assess and *HL*	0.61***	0.31*
Self-assess and SG_{Avg}	0.60**	0.42***
	Game Theory	***No Game Theory***
Self-assess and *HL*	0.43**	0.52***
Self-assess and SG_{Avg}	0.54***	0.37*

Significance level: * p < 0.1, ** p < 0.05, *** p < 0.01.

The former result could be explained by the higher level of complexity of *SG* – where both probability and outcome are different for each lottery in each of the four panels – with respect to *HL*, where the two pairs of outcomes are fixed for all pairwise comparisons. Thus, having some background in

mathematically-related disciplines could be helpful in understanding a risk-elicitation task (in our case, *SG*), although Brañas-Garza et al. (2008) has found no such effect across several risk-elicitation tasks.

Our interpretation is partially supported by a correlation coefficient (between self-assessed risk and the instrument) increasing in the **years of study at the university** for undergraduate students (from first to third year), if considering either SG_{Avg} or *HL* as risk-elicitation instrument (see Table 4.4). Notice that, although quite high, few correlations are statistically significant, due to few observations in each subset of subjects.

Table 4.4: Rank correlations between self-assessed risk and the two tasks, disentangled by the years of study at the university

	First		**Second**		**Third**		**Fourth**		**Fifth**	
	rho	*p-value*	*rho*	*p-value*	*rho*	*p-value*	*rho*	*p-value*	*rho*	*p-value*
Self-assess and SG_{Avg}	0.30	0.160	0.59	0.070	0.62	0.055	0.53	0.140	-0.01	0.980
Self-assess and HL	0.24	0.270	0.79	0.006	0.67	0.030	0.08	0.840	0.57	0.140

Risk-related motivations

A further step is to consider the possibility that there are *several* risk-related motivations that drive subjects' choices among lotteries in *SG*. With this in mind, we split the sample into three categories, according to the three main patterns of choices a subject can show in the four panels of *SG*.

The baseline category is the one comprising subjects whose behavioral pattern across panels is coherent with a Constant Relative Risk Aversion utility function (**CRRA-coherent** subjects). In our sample, subjects showing such compatibility are 14/62: They make *weakly-increasing choices* (in terms of risk-taking) in the four panels. For example, a subject belonging to this category select a number of tickets assigned to a positive outcome of 7/10 in panel 1, of 6/10 in panel 2, of 6/10 in panel 3, and of 2/10 in panel 4. Due to the discreteness of the decision settings, we include in this group also subjects who always made the same choice in the four panels: Although, as we will see below, they could reasonably belong to the other two categories, their behavior do not show incoherence with *CRRA*.

The second category includes subjects with *weakly-decreasing choices* (in terms or risk) in the four panels, hence incoherent with expected-utility maximization. In our sample, subjects showing this behavior are 17/62. We call them **Aspiration-level** subjects. Indeed, it is well known in the literature (Camerer et al., 1997; Diecidue and Van De Ven, 2008) that the concept of *aspiration level* is related to the subject's willingness to reach a particular outcome. In the paper by Camerer et al. (1997) the idea of aspiration level is explained through the cab drivers example: Cab drivers are willing to earn a daily target return, so that they adjust this behavior in order to achieve their goal. Other examples have been proposed in the literature such as farmers who want to prevent themselves from falling below the subsistence level (Lopes, 1987) or investors with the desired target rate of returns to achieve (Payne et al., 1980). In our framework, the idea of aspiration level could be explained by the willingness of our subjects to earn "around a given positive amount". Given the structure of the four panels in *SG*, the risk that one should take to get the "same" positive amount is smaller (the number of winning tickets is higher) the higher the panel number. For example, suppose that a subject wants to earn around €8 in each of the four panels. This is consistent with selecting a number of tickets assigned to this outcome equal to 1/10 in panel 1, equal to 2/10 in panel 2, equal to 4/10 in panel 3, and equal to 6/10 in panel 4.

The residual category is composed by individuals who show non-monotonic choices in the four panels: They "move right and left" across the four panels. We call them **Hedging** subjects. The intuition behind this label is that these subjects might interpret the four panels in *SG* as a *portfolio of contingent assets* (indeed, only one of the four panels is randomly selected for payment), where they can compensate the greater risk taken in some state of the world (e.g., in panel 1 and in panel 3) by choosing less risky assets in the complementary states (e.g., in panel 2 and in panel 4). This intuition is indirectly shared by Crosetto and Filippin (2015): in motivating their between-subject design, they underline how proposing several risky choices on a within-subject base is likely to induce some form of hedging across panels. In our sample, subjects showing this behavior are 31/62.

Both Aspiration-level and Hedging can be viewed as additional risk-related motivations (additional with respect to *CRRA*). Therefore, two behavioral hypotheses can be drawn about behavior in *SG*:

H1) Aspiration-level vs. *CRRA*-coherent. Subjects with a given aspiration level pick a willing-to-win amount in the first panel of lotteries (the one with the lowest payoffs) and then decrease the probability of winning in the next panels, where payoffs are increased, in order to get around this amount. This ends up in a more risk-averse behavior in *SG* than the one disclosed in the same task by *CRRA*-coherent subjects. Indeed, the structure of *SG* contraints choices of an Aspiration-level subject in the four panels not to be too "far away" from one another, i.e. he/she chooses lotteries with close numbers of winning tickets in the four panels (e.g. earning around €5 requires choosing a lottery with 2/10, 3/10, 6/10 and 7/10 tickets assigned to the positive outcome respectively in panel 1, 2, 3 and 4 – see Figure A2 in the Appendix). This ends up in a lower variance of the expected values of the chosen lotteries in the four panels, with respect to *CRRA*-coherent subjects. The latter, given a degree of risk aversion r, when moving from panel i to panel $i+1$, are "free" to choose lotteries with higher expected values, i.e. with number of winning tickets in panel $i+1$ potentially much higher than in panel i (e.g. a *CRRA*-coherent subject with $r = 0.091$ would choose a lottery with 10/10, 2/10, 1/10 and 1/10 tickets assigned to the positive outcome respectively in panel 1, 2, 3 and 4 – see Table A2 in the Appendix).

H2) Hedging vs. *CRRA*-coherent. Subjects who hedge among lottery panels in *SG* are more risk-averse than pure *CRRA*-coherent subjects. Indeed, hedging among contingent assets (lottery panels) introduces additional contraints to the set of lotteries a subject can choose in panel i given the choice made in the other three panels $j \neq i$. For example, a *CRRA*-coherent subject with $r > 0.1$ would choose a lottery with 10/10 winning tickets in panel 1 and with 1/10 winning tickets in panel 4. An Hedging subject with the same r would risk more when stakes are smaller (e.g., by choosing 5/10 winning tickets in panel 1), compensating this riskier choice by risking less when stakes are bigger (e.g., by choosing 5/10 winning tickets

in panel 4). This ultimately leads to a lower variance of the expected values of the chosen lotteries in the four panels, with respect to *CRRA*-coherent subjects.

In order to test H1 and H2 we look at the variance among the expected values in the four chosen lotteries in *SG*, for each subject and for each category of subjects. In this task, the variance for each subject is a measure of the dispersion of the four choices with respect to the mean choice.

We find that both Aspiration-level subjects and Hedging subjects have a lower average variance across panels of lottery expected values (respectively, 4.62 and 6.98) with respect to *CRRA*-coherent subjects (9.50). Both these differences are significant at the 5% level.

As a further round of investigation we perform a Mann-Whitney test by categories on standard deviations of "chosen" expected values in the four panels. Taking *CRRA*-coherent subjects are reference category, we find that the rank of these standard deviations is significantly lower for both Aspiration-level subjects (*p-value*=0.008) and for Hedging subjects (*p-value*=0.000). All this is summarized in the following:

Result 4. *Both Aspiration-level and Hedging subjects disclose in SG a more risk-averse behavior than CRRA-coherent subjects.*

Now we check whether by disentangling the sample according to the three above categories, the correlation between disclosed orderings of risk behavior in the two instruments increases. To this goal, we run another rank correlation test, and we find that the coefficients are higher with respect to the whole sample but still not significant (see Table 4.5).

Table 4.5: Rank correlations among instruments (*HL* and SG_{Avg}), disentangled by category of subjects

	SG			
	Whole Sample	*CRRA*	*Aspiration*	*Hedging*
HL	0.13	0.21	0.22	0.19

Significance level: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The following statement extends Result 2:

Result 5. *If we disentangle by different risk-related motivations, the rank correlation between risk orderings in HL and SG increases. However, the correlation is still not significant.*

Going back to Result 3, we can improve the analysis of the goodness of the two instruments in disclosing subjects' self-assessed general risk aversion, by running the usual Spearman correlation test on each of the above defined categories of risk-related motivations (see Table 4.6). Disentangling by risk-related motivation, we find that *HL* performs on average better than *SG* in disclosing a subject's self-assessed general risk aversion. However, while *CRRA*-coherent subjects show a greater rank correlation with self-reported risk in *HL* than in *SG*, the latter better captures self-assessed risk aversion of Hedging subjects. None of the instruments is able to elicit the self-assessed general risk aversion of Aspiration-level subjects.

Table 4.6: Self-assessed risk, disentangled by category of subjects

	Self-assessed risk			
	Whole Sample	*CRRA*	*Aspiration*	*Hedging*
HL	0.47***	0.80***	0.38	0.45**
SG	0.48***	0.39	0.17	0.66***

Significance level: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The following statement extends Result 3:

Result 6. *When focusing only on subjects whose behavior is coherent with CRRA within an expected utility framework, HL is able to capture 80% of differences in subjects' self-assessed general aversion to risk. SG better captures self-assessed differences for subjects whose behavior is not coherent with CRRA, when they show hedging behavior – hence, additional demand for risk protection – in this task.*

Finally, we focus on the determinants of behavior in *HL*. We use the average choice among the four panels in *SG* and the hypothetical question – together with the other questions on idiosyncratic features in the final questionnaire – as regressors (see Table 4.7).

Table 4.7: Determinants of behavior in *HL* (OLS regression)

	CRRA	***Aspiration***	***Hedging***	***Whole sample***
SG_{Avg}	–1.87*	–0.93	0.30	–0.42
Self-assessment	0.92	2.43**	–0.12	1.04***
Gender	1.16	–1.11	4.05*	0.05
Age	–0.60	–0.27	2.52**	–0.05
Years of study	1.00	0.07	–2.40*	0.05
Study	0.42	–0.59	0.28	0.61
Game Theory	1.00	2.91	–2.56*	–0.44
Constant	3.88	–16.77	–64.62***	–17.50**
Obs.	17	14	31	62

Significance level: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

First of all, the results in Table 4.7 show the strong relation between risk behavior in *HL* and self-assessed aversion to general risk: the latter has a highly-significant impact on risk-averse behavior in *HL*.

Second, the regression analysis clarifies why *SG* is not a good predictor of risk behavior in *HL*: the two instruments lead to opposite behavior for *CRRA*-coherent subjects, this category being the one whose behavior is better predicted by *HL*.

Furthermore, Aspiration-level subjects' behavior in *HL* is driven by their self-reported general aversion to risk. The intuition is that the willingness to take general risks determines a switch line in *HL* that in turn mirrors the specific (expected) outcome these subjects wish to obtain.

Finally, the regression analysis show that idiosyncratic features have an effect on risk behavior in *HL* only for Hedging subjects. The intuition is that the heterogeneity of possible hedging strategies (several possible non-monotonic patterns in *SG*) might hide the interplay of idiosyncratic features, that ultimately impact on behavior in *HL*.

We conclude with some technical remarks. As far as we noted that the t-statistics for the coefficients are only marginally significant but with an overall F strongly significant, we conducted analyses to check for the possibility of multicollinearity. What we found is that the cross-correlations are low except for the variables Age and Years of Study (high correlations between pairs of coefficients would have indicated possible multicollinearity

problems). As a further round of investigation, we run multicollinearity tests and we looked at the condition number, that is actually high (60.712), as well as condition indexes. On the contrary, Variable Inflation Factors are small. The diagnostics widely disagree. This is neither a surprising finding nor a problem from the point of view of the results. In fact, even extreme multicollinearity (and this is not the case under consideration here) does not violate the OLS assumptions: OLS estimates are still unbiased and BLUE (Best Linear Unbiased Estimators). We actually tried many different specifications of our regression model on the same data set. None of these changes produced significant improvements, suggesting that multicollinearity is not a relevant problem to be considered here.[13]

The following statement summarizes the main findings about the determinants of behavior in *HL*:

Result 7. *Self-assessed risk appears to be a relevant determinant of risk-related choices in HL, especially for Aspiration-level subjects. For Hedging subjects, choices in HL are not explained by either SG or self-assessed risk; they are rather driven by idiosyncratic features.*

4.6 Conclusion

In this paper, we deep delve into a well-established result of the literature on risk elicitation: Making use of different experimental methods leads to different results in elicited risk preferences among subjects. To this end, we compare two Multiple Price List Design methods, one based on a single-choice setting (*HL*) and the other on a multiple-choice one (*SG*).

As a first step, we make use of usual non-parametric statistical tools to check whether subjects facing our different tasks at least maintain the same ordering in their risk-related lottery choices. Apparently, what we find confirms the common result of independence among instruments. As a matter of fact, the rank correlation between the two instruments turns out

[13] To account for the ordered nature of our dependent variable, we have also estimated an Ordered Logit Model: the results are qualitatively unchanged and available upon request.

to be negligible and not significant.

Our analysis goes beyond by making use of a self-assessed measure of subjects' risk preferences, to check whether our instruments are capable to measure differences in self-reported attitudes towards risk.

What we find is that for both risk-elicitation procedures, risk preferences disclosed by subjects' choices are significantly correlated with their self-reported risk. This result is even stronger when we run a by-group analysis on different idiosyncratic controls.

Furthermore, and more importantly, we check whether other subjects' risk-related motivations could explain the correlation between elicited risk behavior through monetary-incentivized methods and self-assessed risk.

Since a multiple-choice risk elicitation method is available (SG), we use it so as to disentangle subjects according to three risk-related behaviors: the baseline behavioral category comprising $CRRA$-coherent individuals, a group of *Aspiration-level* subjects, and a last category of *Hedging* subjects.

What is found is that in SG both Aspiration-level and Hedging subjects make on average less risky choices than $CRRA$-coherent subjects. This confirms the intuition that both these categories hide an additional risk-related motivation (additional with respect to the curvature of the uni-parametric utility function), that cannot be disentangled by only looking at behavior in HL.

It is not surprising that if we exclude the two above categories and we only focus on subjects whose behavior in SG is coherent with $CRRA$, HL is able to capture 80% of differences in subjects' self-assessed general aversion to risk. A regression analysis confirms that self-assessed risk is a relevant determinant of risk-related choices in HL.

This result is even more striking when considering that it was obtained in a within-subject design. Indeed, as Crosetto and Filippin (2015) correctly notice, proposing several risky choices on a within-subject base is likely to induce some form of hedging across tasks by non-risk-averse subjects. This could determine a negative correlation across tasks. Therefore, the low correlation between the behavior in different tasks could in part be an artifact of the design. We have shown that once this hedging motivation is set aside

(i.e. only *CRRA*-coherent subjects are considered), despite no correlation between *HL* and *SG*, an extremely high correlation between the risk behavior in the former method and self-reported risk attitude emerges. Thus, the positive results by Vieider et al. (2015) might hold even stronger if we account for heterogeneity stemming from more complex behavioral patterns like aspiration levels and hedging.

This result is relevant for experimental economists who wish to account for their subjects' risk attitudes as a determinant of behavior, being *HL* the experimental method mainly used in the last decade to elicit risk attitudes in the laboratory and in the field.

Appendix

Figure A1: Task 2: Variant of Holt and Laury (2002)

	LOTTERY **A**	LOTTERY **B**
L1	If the drawn ticket is no. **1**, you win ***12.00*** euros; otherwise, if the drawn ticket is between **2** and **20**, you win ***10.00*** euros.	If the drawn ticket is no. **1**, you win ***22.00*** euros; otherwise, if the drawn ticket is between **2** and **20**, you win ***0.50*** euros.
L2	If the drawn ticket is between **1** and **2**, you win ***12.00*** euros; otherwise, if the drawn ticket is between **3** and **20**, you win ***10.00*** euros.	If the drawn ticket is between **1** and **2**, you win ***22.00*** euros; otherwise, if the drawn ticket is between **3** and **20**, you win ***0.50*** euros.
L3	If the drawn ticket is between **1** and **3**, you win ***12.00*** euros; otherwise, if the drawn ticket is between **4** and **20**, you win ***10.00*** euros.	If the drawn ticket is between **1** and **3**, you win ***22.00*** euros; otherwise, if the drawn ticket is between **4** and **20**, you win ***0.50*** euros.
L4	If the drawn ticket is between **1** and **4**, you win ***12.00*** euros; otherwise, if the drawn ticket is between **5** and **20**, you win ***10.00*** euros.	If the drawn ticket is between **1** and **4**, you win ***22.00*** euros; otherwise, if the drawn ticket is between **5** and **20**, you win ***0.50*** euros.
L5	If the drawn ticket is between **1** and **5**, you win ***12.00*** euros; otherwise, if the drawn ticket is between **6** and **20**, you win ***10.00*** euros.	If the drawn ticket is between **1** and **5**, you win ***22.00*** euros; otherwise, if the drawn ticket is between **6** and **20**, you win ***0.50*** euros.
L6	If the drawn ticket is between **1** and **6**, you win ***12.00*** euros; otherwise, if the drawn ticket is between **7** and **20**, you win ***10.00*** euros.	If the drawn ticket is between **1** and **6**, you win ***22.00*** euros; otherwise, if the drawn ticket is between **7** and **20**, you win ***0.50*** euros.
L7	If the drawn ticket is between **1** and **7**, you win ***12.00*** euros; otherwise, if the drawn ticket is between **8** and **20**, you win ***10.00*** euros.	If the drawn ticket is between **1** and **7**, you win ***22.00*** euros; otherwise, if the drawn ticket is between **8** and **20**, you win ***0.50*** euros.
L8	If the drawn ticket is between **1** and **8**, you win ***12.00*** euros; otherwise, if the drawn ticket is between **9** and **20**, you win ***10.00*** euros.	If the drawn ticket is between **1** and **8**, you win ***22.00*** euros; otherwise, if the drawn ticket is between **9** and **20**, you win ***0.50*** euros.
L9	If the drawn ticket is between **1** and **9**, you win ***12.00*** euros; otherwise, if the drawn ticket is between **10** and **20**, you win ***10.00*** euros.	If the drawn ticket is between **1** and **9**, you win ***22.00*** euros; otherwise, if the drawn ticket is between **10** and **20**, you win ***0.50*** euros.
L10	If the drawn ticket is between **1** and **10**, you win ***12.00*** euros; otherwise, if the drawn ticket is between **11** and **20**, you win ***10.00*** euros.	If the drawn ticket is between **1** and **10**, you win ***22.00*** euros; otherwise, if the drawn ticket is between **11** and **20**, you win ***0.50*** euros.
L11	If the drawn ticket is between **1** and **11**, you win ***12.00*** euros; otherwise, if the drawn ticket is between **12** and **20**, you win ***10.00*** euros.	If the drawn ticket is between **1** and **11**, you win ***22.00*** euros; otherwise, if the drawn ticket is between **12** and **20**, you win ***0.50*** euros.
L12	If the drawn ticket is between **1** and **12**, you win ***12.00*** euros; otherwise, if the drawn ticket is between **13** and **20**, you win ***10.00*** euros.	If the drawn ticket is between **1** and **12**, you win ***22.00*** euros; otherwise, if the drawn ticket is between **13** and **20**, you win ***0.50*** euros.
L13	If the drawn ticket is between **1** and **13**, you win ***12.00*** euros; otherwise, if the drawn ticket is between **14** and **20**, you win ***10.00*** euros.	If the drawn ticket is between **1** and **13**, you win ***22.00*** euros; otherwise, if the drawn ticket is between **14** and **20**, you win ***0.50*** euros.
L14	If the drawn ticket is between **1** and **14**, you win ***12.00*** euros; otherwise, if the drawn ticket is between **15** and **20**, you win ***10.00*** euros.	If the drawn ticket is between **1** and **14**, you win ***22.00*** euros; otherwise, if the drawn ticket is between **15** and **20**, you win ***0.50*** euros.
L15	If the drawn ticket is between **1** and **15**, you win ***12.00*** euros; otherwise, if the drawn ticket is between **16** and **20**, you win ***10.00*** euros.	If the drawn ticket is between **1** and **15**, you win ***22.00*** euros; otherwise, if the drawn ticket is between **16** and **20**, you win ***0.50*** euros.
L16	If the drawn ticket is between **1** and **16**, you win ***12.00*** euros; otherwise, if the drawn ticket is between **17** and **20**, you win ***10.00*** euros.	If the drawn ticket is between **1** and **16**, you win ***22.00*** euros; otherwise, if the drawn ticket is between **17** and **20**, you win ***0.50*** euros.
L17	If the drawn ticket is between **1** and **17**, you win ***12.00*** euros; otherwise, if the drawn ticket is between **18** and **20**, you win ***10.00*** euros.	If the drawn ticket is between **1** and **17**, you win ***22.00*** euros; otherwise, if the drawn ticket is between **18** and **20**, you win ***0.50*** euros.
L18	If the drawn ticket is between **1** and **18**, you win ***12.00*** euros; otherwise, if the drawn ticket is between **19** and **20**, you win ***10.00*** euros.	If the drawn ticket is between **1** and **18**, you win ***22.00*** euros; otherwise, if the drawn ticket is between **19** and **20**, you win ***0.50*** euros.
L19	If the drawn ticket is between **1** and **19**, you win ***12.00*** euros; otherwise, if the drawn ticket is no. **20**, you win ***10.00*** euros.	If the drawn ticket is between **1** and **19**, you win ***22.00*** euros; otherwise, if the drawn ticket is no. **20**, you win ***0.50*** euros.

Figure A2: Task 2: Sabater-Grande and Georgantzis (2002)

PANEL 1

Prob. of Winning	**100%**	**90%**	**80%**	**70%**	**60%**	**50%**	**40%**	**30%**	**20%**	**10%**
Euros	**1**	**1.12**	**1.27**	**1.47**	**1.73**	**2.10**	**2.65**	**3.56**	**5.40**	**10.90**
I prefer										

PANEL 2

Prob. of Winning	**100%**	**90%**	**80%**	**70%**	**60%**	**50%**	**40%**	**30%**	**20%**	**10%**
Euros	**1**	**1.20**	**1.50**	**1.90**	**2.30**	**3**	**4**	**5.70**	**9**	**19**
I prefer										

PANEL 3

Prob. of Winning	**100%**	**90%**	**80%**	**70%**	**60%**	**50%**	**40%**	**30%**	**20%**	**10%**
Euros	**1**	**1.66**	**2.50**	**3.57**	**5**	**7**	**10**	**15**	**25**	**55**
I prefer										

PANEL 4

Prob. of Winning	**100%**	**90%**	**80%**	**70%**	**60%**	**50%**	**40%**	**30%**	**20%**	**10%**
Euros	**1**	**2.20**	**3.80**	**5.70**	**8.30**	**12**	**17.50**	**26.70**	**45**	**100**
I prefer										

Table A1: Elicited r for *HL*

Switch Line	Elicited Degree of Risk Aversion
$L1$	$-\infty < r \leq -2.863$
$L2$	$-2.863 \leq r \leq -2.010$
$L3$	$-2.010 \leq r \leq -1.512$
$L4$	$-1.512 \leq r \leq -1.156$
$L5$	$-1.156 \leq r \leq -0.878$
$L6$	$-0.878 \leq r \leq -0.648$
$L7$	$-0.648 \leq r \leq -0.450$
$L8$	$-0.450 \leq r \leq -0.273$
$L9$	$-0.273 \leq r \leq -0.112$
$L10$	$-0.112 \leq r \leq 0.038$
$L11$	$0.038 \leq r \leq 0.180$
$L12$	$0.180 \leq r \leq 0.317$
$L13$	$0.317 \leq r \leq 0.454$
$L14$	$0.454 \leq r \leq 0.592$
$L15$	$0.592 \leq r \leq 0.736$
$L16$	$0.736 \leq r \leq 0.891$
$L17$	$0.891 \leq r \leq 1.068$
$L18$	$1.068 \leq r \leq 1.287$
$L19$	$1.287 \leq r \leq 1.613$
$L20$	$1.613 \leq r < \quad +\infty$

Table A2: Elicited r for *SGG*

Panel 1		Panel 2		Panel 3		Panel 4	
Chosen p	Elicited r	Chosen p	Elicited r	Chosen p	Elicited r	Chosen p	Elicited r
1	0.091	1	0.500	1	0.833	1	0.909
0.9	0.082	0.9	0.450	0.9	0.750	0.9	0.818
0.8	0.073	0.8	0.400	0.8	0.667	0.8	0.727
0.7	0.064	0.7	0.350	0.7	0.583	0.7	0.636
0.6	0.055	0.6	0.300	0.6	0.500	0.6	0.545
0.5	0.045	0.5	0.250	0.5	0.417	0.5	0.455
0.4	0.036	0.4	0.200	0.4	0.333	0.4	0.364
0.3	0.027	0.3	0.150	0.3	0.250	0.3	0.273
0.2	0.018	0.2	0.100	0.2	0.167	0.2	0.182
0.1	0.009	0.1	0.050	0.1	0.083	0.1	0.091

Bibliography

Andersen, S., Harrison, G. W., Lau, M. I., and Rutström, E. E. (2006). Elicitation using multiple price list formats. *Experimental Economics*, 9:383–405.

Attanasi, G., Casoria, F., Centorrino, S., and Urso, G. (2013). Cultural investment, local development and instantaneous social capital: A case study of a gathering festival in the south of italy. *Journal of Socio-Economics*, 47:228–247.

Attanasi, G., Corazzini, L., Georgantzís, N., and Passarelli, F. (2014a). Risk aversion, overconfidence and private information as determinants of majority thresholds. *Pacific Economic Review*, 19:355–386.

Attanasi, G., Gollier, C., Montesano, A., and Pace, N. (2014b). Eliciting ambiguity aversion in unknown and in compound lotteries: A smooth ambiguity model experimental study. *Theory and Decision*, 77:485–530.

Becker, G. M., DeGroot, M. H., and Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*, 9:226–232.

Bernasconi, M., Corazzini, L., and Seri, R. (2014). Reference dependent

preferences, hedonic adaptation and tax evasion: Does the tax burden matter? *Journal of Economic Psychology*, 40:103–118.

Brañas-Garza, P., Guillen, P., and del Paso, R. L. (2008). Math skills and risk attitudes. *Economics Letters*, 99:332–336.

Butler, D. J. and Loomes, G. C. (2007). Imprecision as an account of the preference reversal phenomenon. *American Economic Review*, 97:277–297.

Camerer, C., Babcock, L., Loewenstein, G., and Thaler, R. (1997). Labor supply of new york city cabdrivers: One day at a time. *Quarterly Journal of Economics*, 112:407–441.

Charness, G. and Gneezy, U. (2010). Portfolio choice and risk attitudes: An experiment. *Economic Inquiry*, 48:133–146.

Charness, G., Gneezy, U., and Imas, A. (2013). Experimental methods: Eliciting risk preferences. *Journal of Economic Behavior and Organization*, 87:43–51.

Charness, G. and Viceisza, A. (2015). Three risk-elicitation methods in the field: Evidence from rural senegal. *Review of Behavioral Economics*, forthcoming.

Cox, J. C. and Harrison, G. W. (2008). *Risk aversion in experiments*. Emerald: Research in Experimental Economics, Vol. 12.

Cox, J. C., Roberson, B., and Smith, V. L. (1982). Theory and behavior of single object auctions. *Research in Experimental Economics*, 2:1–43.

Cox, J. C., Sadiraj, V., and Schmidt, U. (2014). Paradoxes and mechanisms for choice under risk. *Experimental Economics*, 18:215–250.

Cox, J. C., Smith, V. L., and Walker, J. M. (1985). Experimental development of sealed-bid auction theory; calibrating controls for risk aversion. *American Economic Review*, pages 160–165.

Cox, J. C., Smith, V. L., and Walker, J. M. (1988). Theory and individual behavior of first-price auctions. *Journal of Risk and Uncertainty*, 1:61–99.

Crosetto, P. and Filippin, A. (2013). The "bomb" risk elicitation task. *Journal of Risk and Uncertainty*, 47:31–65.

Crosetto, P. and Filippin, A. (2015). A theoretical and experimental appraisal of five risk elicitation methods. *Experimental Economics*, forthcoming.

Dave, C., Eckel, C. C., Johnson, C. A., and Rojas, C. (2010). Eliciting risk preferences: When is simple better? *Journal of Risk and Uncertainty*, 41:219–243.

Diecidue, E. and Van De Ven, J. (2008). Aspiration level, probability of success and failure, and expected utility. *International Economic Review*, 49:683–700.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9:522–550.

Eckel, C. C. and Grossman, P. J. (2008). Forecasting risk attitudes: An experimental study using actual and forecast gamble choices. *Journal of Economic Behavior and Organization*, 68:1–17.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10:171–178.

García Gallego, A., Georgantzís, N., Jaramillo Gutiérrez, A., and Parravano, M. (2012). The lottery-panel task for bi-dimensional parameter-free elicitation of risk attitudes. *Revista Internacional de Sociologia*, 70:53–72.

García-Gallego, A., Georgantzís, N., Navarro-Martínez, D., and Sabater-Grande, G. (2011). The stochastic component in choice and regression to the mean. *Theory and Decision*, 71:251–267.

Georgantzís, N. and Navarro-Martínez, D. (2010). Understanding the wta–wtp gap: Attitudes, feelings, uncertainty and personality. *Journal of Economic Psychology*, 31:895–907.

Guiso, L. and Paiella, M. (2008). Risk aversion, wealth, and background risk. *Journal of the European Economic association*, 6:1109–1150.

Harrison, G. W., Johnson, E., McInnes, M. M., and Rutström, E. E. (2005). Risk aversion and incentive effects: Comment. *American Economic Review*, 95:897–901.

Harrison, G. W. and Rutström, E. E. (2008). Risk aversion in the laboratory. In Harrison, G. W. and Cox, J., editors, *Risk aversion in experiments*, pages 41–196. Emerald: Research in Experimental Economics, Vol. 12.

Harrison, G. W. and Rutström, E. E. (2009). Expected utility theory and prospect theory: One wedding and a decent funeral. *Experimental Economics*, 12:133–158.

Holt, C. A. and Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, 92:1644–1655.

Holt, C. A. and Laury, S. K. (2005). Risk aversion and incentive effects: New data without order effects. *American Economic Review*, 95:902–904.

Isaac, R. M. and James, D. (2000). Just who are you calling risk averse? *Journal of Risk and Uncertainty*, 20:177–187.

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., Strong, D. R., and Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: the balloon analogue risk task (bart). *Journal of Experimental Psychology: Applied*, 8:75–84.

Lévy-Garboua, L., Maafi, H., Masclet, D., and Terracol, A. (2012). Risk aversion and framing effects. *Experimental Economics*, 15:128–144.

Lopes, L. L. (1987). Between hope and fear: The psychology of risk. *Advances in Experimental Social Psychology*, 20:255–295.

Miller, L., Meyer, D. E., and Lanzetta, J. T. (1969). Choice among equal expected value alternatives: Sequential effects of winning probability level on risk preferences. *Journal of Experimental Psychology*, 79:419–423.

Payne, J. W., Laughhunn, D. J., and Crum, R. (1980). Translation of gambles and aspiration level effects in risky choice behavior. *Management Science*, 26:1039–1060.

Roth, A. E. and Malouf, M. W. (1979). Game-theoretic models and the role of information in bargaining. *Psychological Review*, 86:574.

Sabater-Grande, G. and Georgantzis, N. (2002). Accounting for risk aversion in repeated prisoners' dilemma games: An experimental test. *Journal of Economic Behavior and Organization*, 48:37–50.

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

Wakker, P. and Deneffe, D. (1996). Eliciting von neumann-morgenstern utilities when probabilities are distorted or unknown. *Management Science*, pages 1131–1150.

Wakker, P. P. (2010). *Prospect theory: For risk and ambiguity*. Cambridge University Press.

Zuckerman, M. (1994). *Behavioral expressions and biosocial bases of sensation seeking*. Cambridge university press.

Acknowledgments

I would like to express my special thanks to my supervisors Professor Dr. Luca Stanca and Professor Dr. Giuseppe Attanasi. Thank you for encouraging me even during tough times in the Ph.D. pursuit. I am thankful for the excellent example you have provided as successful economists and as professors (actually as masters).

Grazie ai miei amici, quelli incontrati durante il percorso di dottorato e quelli incontrati lontano nel tempo. Senza di voi non sarebbe stato cosí divertente. Un grazie speciale alla mia famiglia. Non so trovare parole adeguate per dirvi quanto vi sia riconoscente. Grazie per la pazienza che avete avuto. Senza il vostro sostegno costante non sarei mai arrivata fin qui. Grazie per aver saputo sognare con me. Grazie anche alla mia nuova famiglia adottiva. Sono certa che questa fatica l'abbiate in qualche modo vissuta anche voi. Ultimo, ma non ultimo, un grazie speciale a mio marito, Leonardo. Tu hai creduto più di me, hai sperato più di me e sei sempre stato il mio sostegno quando non sapevo più credere.

This thesis is also yours.